The Evolution of Casemix Measurement Using Diagnosis Related Groups (DRGs)

Richard F. Averill, M.S., John H. Muldoon, M.H.A., James C. Vertrees, Ph.D., Norbert I. Goldfield, M.D, Robert L. Mullin, M.D., Elizabeth C. Fineran, M.S., Mona Z. Zhang, B.S., Barbara Steinbeck, ART, Thelma Grant, RRA

The Diagnosis Related Groups (DRGs) have been evolving for more than twenty years. The use of DRGs in a wide variety of applications has resulted in the development of five distinct versions of the DRGs: Medicare DRGs, Refined DRGs, Severity DRGs, All Patient DRGs and All Patient Refined DRGs. This article provides a description of the structure and major features of each DRG system and a comparison of the statistical performance of each system. Using a 1993 database of 675 acute care hospitals plus a twenty percent sample of 40 children's hospitals, the ability of the DRG systems to predict cost and mortality at the patient level was measured in several ways including the R^2 *for cost and* R^2 *and c-statistic for mortality. A hospital level* R^2 *for cost and mortality was also computed for each system using a multiple regression that included hospital characteristics. A payment simulation was performed and the payment redistribution impact of each DRG system was evaluated. Results are presented overall and across a broad range of patient and hospital characteristics. The results show substantial variations across all measures with the APR-DRGs, in general, having the best performance.*

This study was performed by 3M Health Information Systems and the National Association of Children's Hospitals and Related Institutions. Address correspondence and requests for reprints to Richard F. Averill, M.S., Research Manager, 3M Health Information Systems, 100 Barnes Road, Wallingford, CT 06492. James C. Vertrees, Ph.D., Norbert I. Goldfield, M.D., Robert L. Mullin, M.D., Elizabeth C. Fineran, M.S., Mona Z. Zhang, B.S., Barbara Steinbeck, ART, Thelma Grant, RRA, are with 3M Health Information Systems, 3M Health Care. John H. Muldoon, M.H.A. is Vice President, National Association of Children's Hospitals and Related Institutions, 401 Wythe Street, Alexandria, Virginia 22314. The opinions expressed are those of the authors and do not necessarily reflect the views or policy positions of 3M Health Informations Systems or the National Association of Children's Hospitals and Related Institutions.

INTRODUCTION

The Diagnosis Related Groups (DRGs) are a patient classification scheme which was originally developed as a means of relating the type of patients a hospital treats (i.e., its casemix) to the costs incurred by hospitals. The design and development of the DRGs began in the late sixties at Yale University (Fetter, et al, 1980). The initial motivation for developing the DRGs was to create an effective framework for monitoring the utilization of services in a hospital setting. The first large-scale application of the DRGs was in the late 1970's in the State of New Jersey. The New Jersey State Department of Health used DRGs as the basis of a prospective payment system (PPS) in which hospitals were paid a fixed DRG specific amount for each patient treated. In 1983 Congress enacted a DRG based PPS for all Medicare patients. Subsequent to the enactment of the Medicare PPS, a number of states and large payors implemented DRG based hospital PPS for non Medicare patients. In addition, DRGs have been used as the basis of global budget allocation and payment in several countries in Western and Eastern Europe as well as Australia.

The evolution of the DRGs and their use as the basic unit of payment in Medicare's hospital reimbursement system represent a recognition of the fundamental role which a hospital's casemix plays in determining its costs. Although the concept of casemix may appear straightforward, clinicians, administrators and regulators have

attached different meanings to this concept. Casemix has been used to refer to an interrelated but distinct set of patient attributes which include severity of illness, risk of dying, prognosis, treatment difficulty, need for intervention and resource intensity. When clinicians use the notion of casemix, they typically are referring to one or more aspects of clinical complexity. For clinicians, a complex casemix refers to greater severity of illness, greater risk of mortality, greater treatment difficulty, poorer prognoses and/or a greater need for intervention. Thus, from a clinical perspective, casemix refers to the condition of the patients treated and the treatment difficulty associated with providing care. On the other hand, administrators and regulators usually use the concept of casemix to indicate that the patients treated require more resources which results in a higher cost of providing care. Thus, from an administrative or regulatory perspective, casemix refers to the resource intensity demands that patients place on an institution. While the two interpretations of casemix are often closely related, they can be very different for certain kinds of patients. For example, while terminal cancer patients are severely ill and have a poor prognosis, they require few hospital resources beyond basic nursing care.

The original purpose of the DRGs was to relate the casemix of a hospital to the resource demands and associated costs experienced by the hospital. Thus, DRGs focused exclusively on resource intensity. Therefore, a hospital having a more complex casemix from a DRG perspective meant that the hospital treated patients who required more hospital resources, but not necessarily that the hospital treated patients having a greater severity of illness, a greater risk of dying, a greater treatment difficulty, poorer prognoses or a greater need for intervention. As the health care industry has evolved there has been increased demand for a patient classification system that can be used for applications beyond resource use, cost and payment. In particular, a patient classification system is needed to

- Compare hospitals across a wide range of resource and outcome measures
- Evaluate differences in inpatient mortality rates
- Implement and support critical pathways
- Facilitate continuous quality improvement projects
- Support internal management and planning systems
- Manage capitated payment arrangements

For a DRG system to meet these needs, the objective of the DRG system needed to be expanded in scope. In order to meet the requirements of these diverse applications five distinct DRG systems have been developed. This article will describe and contrast the five DRG systems and evaluate the performance of each system using a 1993 national sample of patient data that includes 675 general acute care hospitals and a 20 percent sample of patients from 40 freestanding acute children's hospitals.

DRG SYSTEMS

DRG technology has experienced an evolutionary development process in which the later generations of DRG systems have incorporated the improvements made by earlier generations.The five DRG systems that have been developed are:

- Medicare DRGs
- Refined DRGs (RDRGs)
- All Patient DRGs (AP-DRGs)
- Severity DRGs (SDRGs)
- All Patient Refined DRGs (APR-DRGs)

Each of these DRG systems were created to address specific limitations in the original DRGs

Medicare DRGs

The initial DRG definitions developed at Yale were intended to describe all the types of patients seen in an acute care hospital. Thus, the DRGs encompassed both the elderly patient population as well as the newborn, pediatric and adult populations. However, with the implementation of the Medicare prospective payment system in October, 1983, the responsibility for the maintenance and modification of the DRG definitions became the responsibility of the Health Care Financing Administration (HCFA). HCFA updates the Medicare DRGs on an annual basis. The focus of all DRG modifications instituted by HCFA subsequent to 1983 has been on problems relating primarily to the elderly population. Version 12.0 of the Medicare DRGs has 492 DRGs and was effective October 1, 1995 and was used in this analysis (Averill, Mullin et al., 1995).

Refined DRGs (RDRGs)

During the mid 1980s the Health Care Financing Administration funded a project at Yale University to revise the use of complications and comorbidities (CCs) in the Medicare DRGs (Freeman, Fetter, et al 1995). In the Medicare DRGs a secondary diagnosis is considered a CC if it causes a significant increase in hospital resource use. For certain types of patients, a different Medicare DRG is assigned depending on whether or not a CC is present. The Yale project mapped all secondary diagnoses that were considered a CC in the Medicare DRGs into 136 secondary diagnosis groups each of which was assigned a CC complexity level. For surgical patients each secondary diagnosis group was assigned to one of four CC complexity levels, (non-CC, moderate CC, major CC and catastrophic CC). For medical patients each secondary diagnosis group was assigned to one of three CC complexity levels (non-CC, moderate or major CC and catastrophic CC). All age splits and CC splits in the Medicare DRGs were eliminated and replaced by the four subgroups for surgical patients, or the three subgroups for medical patients. The Yale project did not reevaluate the categorization of secondary diagnosis as a CC versus a non-CC. Only the diagnoses on the standard Medicare CC list were used to create the moderate, major and catastrophic subgroups. The DRG system developed by the Yale project is referred to as Refined DRGs or RDRGs.

In the RDRGs, all secondary diagnoses in a secondary diagnosis group are assigned the same level and a patient is assigned to the subgroup corresponding to the highest level secondary diagnosis. Thus, a surgical patient with two moderate CCs and one major CC is assigned to the major CC subgroup. The number of secondary diagnoses has no effect on the subgroup assigned to the patient (i.e., multiple secondary diagnoses at one level do not cause a patient to be assigned to a higher subgroup). Thus, although a surgical patient may have four moderate CCs present, the patient is still assigned to the moderate CC subgroup.

In general, the categorization of the secondary diagnosis groups as moderate, major or catastrophic was independent of the DRG to which the patient was assigned. However, in a limited number of cases the standard level for a secondary diagnosis group was modified for a specific MDC or RDRG.

For medical patients in each MDC a

separate DRG was created for early deaths primarily as a means of identifying patients admitted for terminal care. Early death DRGs were not created for surgical patients since patients admitted for terminal care would normally not have surgery performed. Early deaths were defined as patients who died within two days of admission. Thus, length of stay as well as death was used to assign the RDRG.

There is no central source for RDRGs and several vendors produce versions of RDRGs. The RDRGs used in this analysis incorporated all changes in Medicare DRGs through version 10.0. There are a total of 1170 RDRGs in Version 10.0 of the RDRGs. There are virtually no differences between version 10.0 and 12.0 of the Medicare DRGs and, therefore, the use of RDRGs updated only through version 10.0 of the Medicare DRGs has no implications for the study results.

The version of RDRGs used for this analysis was produced by Information Strategies of McGraw Hill, Inc. It is different from the original Yale RDRGs in three ways. First, it retained pediatric age splits for ten of the forty-three Medicare DRGs that have pediatric age spits. Second, instead of attempting to implement complexity subclasses for mental health and substance abuse DRGs, it used pediatric and adult age subgroups for the three mental health DRGs with significant case volume. Third, it introduced some further refinements to the complexity subclasses based upon data analysis performed subsequent to the completion of the Yale RDRG study project. Thus, this version of RDRGs follows the original Yale RDRG model with some updates that its developers felt were improvements along with technical updates to incorporate annual changes to the Medicare DRGs.

All Patient DRGs (AP-DRGs)

In 1987, the state of New York passed legislation instituting a DRG-based prospective payment system for all non-Medicare patients. The legislation included a requirement that the New York State Department of Health (NYDH) evaluate the applicability of the Medicare DRGs to a non-Medicare population. In particular, the legislation required that the DRGs be evaluated with respect to neonates and patients with Human Immunodeficiency Virus (HIV) infections. The evaluation concluded that the Medicare DRGs were not adequate for a non-Medicare population. NYDH entered into an agreement with 3M Health Information Systems (3M HIS) to research and develop all necessary DRG modifications. The DRG definitions developed by NYDH and 3M HIS are referred to as the All Patient DRGs (AP-DRGs).

During the mid 1980's extensive research had been performed independently by the National Association of Children's Hospitals and Related Institutions (NACHRI) on alternative approaches to improving the DRG categories for neonates and other pediatric patients (Berry, et al., 1986). The system developed by NACHRI was called the Pediatric Modified Diagnosis Related Groups or PM-DRGs. The PM-DRGs created many additional DRGs specifically for neonatal and pediatric patients. As part of the New York DRG evaluation effort, NYDH and 3M HIS examined the NACHRI neonatal definitions and adopted a modified version of them. For example, the neonatal PM-DRG modifications introduced birth weight as a classification variable. In the AP-DRGs six birth weight categories are used as the primary variable in formulating the neonatal AP-DRGs.

In addition to the changes for the neonatal AP-DRGs, MDC 24 was created for HIV infection patients. Assignment to MDC 24 is based on a principal diagnosis of an HIV infection, or a principal diagnosis of an HIV related complication combined with a secondary diagnosis of an HIV infection (e.g. principal diagnosis of pneumocystosis and a secondary diagnosis of an HIV infection). The initial release of the AP-DRGs in January of 1988 consisted of the addition of MDC 24 and the restructuring of the neonatal MDC.

Subsequent updates have made additional enhancements to the AP-DRGs. For example, as demonstrated by the RDRGs some complications and comorbidities (CC) have a greater impact on hospital resource use than other CCs. The AP-DRGs designate a subset of the CCs as major CCs. These major CCs are similar to the catastrophic CCs in the RDRGs. In order to avoid significantly increasing the number of DRGs, major CC AP-DRGs for surgical patients within an MDC and major CC AP-DRGs for medical patients within an MDC were formed for some MDCs. In total, 60 major CC AP-DRGs were created. As part of the development of the major CC AP-DRGs, a limited reevaluation of the CC list was performed.

The AP-DRGs introduced many other changes to the Medicare DRGs. Some of these primarily affect pediatric patients while others affect patients of all ages. The pediatric modifications include some of the recommendations originally developed by NACHRI as well as other significant modifications. MDC 25 was added for patients with multiple trauma. In addition, significant modifications have been made for transplants, long term mechanical ventilation patients, cystic fibrosis, nutritional disorders, high risk obstetric care, acute leukemia, hemophilia and sickle cell anemia.

Some of the DRG modifications originally developed in the AP-DRGs have subsequently been adopted in the Medicare DRGs. For example, in Version 8.0 of the Medicare DRGs, an HIV infection MDC was added. However, the Medicare HIV infection MDC consists of three DRGs and does not discriminate among HIV infection patients at the level of detail contained in the AP-DRGs.

The AP-DRGs have been updated in January of every year since 1988. Version 12.0 of the AP-DRGs was effective January 1, 1995, and was used in this analysis (Averill, Mullin, et al 1995). There are a total of 641 AP-DRGs in Version 12.0 of the AP-DRGs.

Severity DRGs (SDRGs)

In 1993, HCFA initiated a reevaluation of the use of complications and comorbidities within the Medicare DRGs (Federal Register, 1994). The DRG system developed by this reevaluation is referred to as the severity DRGs or SDRGs (sometimes also referred to as Severity Refined DRGs or SR-DRGs). The reevaluation excluded the DRGs associated with pregnancy, newborns and pediatric patients. The major CC list from the AP-DRGs was used to identify an initial list of major CCs. Using Medicare data the categorization of each secondary diagnosis as a non-CC, non major CC or a major CC was reevaluated. The end result was that 111 diagnoses that were non-CCs in the Medicare DRGs were made a CC, 220 diagnoses that were a CC were made a non-CC and 395 CCs were considered a major CC. This evaluation differed from the RDRGs in that both non-CCs and CCs were evaluated and the evaluation was performed independently for each individual diagnosis as opposed to aggregate secondary diagnosis groups.

All CC splits in the Medicare DRGs were

eliminated plus an additional 24 Medicare DRGs were merged together. The resulting base DRGs were then subdivided into three, two or no subgroups based on an analysis of Medicare data. The result was 84 DRGs with no subgroups, 124 with two subgroups and 85 with three subgroups. This plus the 63 DRGs not evaluated resulted in a total of 652 SDRGs.

In SDRGs, a patient is assigned to a subgroup corresponding to the highest level secondary diagnosis. Like RDRGs, multiple secondary diagnoses at a one level do not cause a patient to be assigned to a higher subgroup. The categorization of a diagnosis as non-CC, non major CC or major CC is uniform across the SDRGs and there are no modifications for specific SDRGs. HCFA published the SDRGs in 1994, but did not establish an implementation date. The SDRGs have not been updated by HCFA since the original 1994 release.

All Patient Refined DRGs (APR-DRGs)

The All Patient Refined DRGs (APR-DRGs) further refine the basic AP-DRG structure by adding four subgroups. In contrast to the SDRGs and RDRGs, the initial version of the APR-DRGs used the AP-DRGs as the base DRGs instead of the Medicare DRGs. The only exception was the neonatal MDC for which an entirely new set of base DRGs was developed. In subsequent updates of the APR-DRGs the base AP-DRGs have been substantially modified. All age, CC and major CC distinctions in the AP-DRGs were eliminated and replaced by two sets of four subgroups. One set of subgroups addresses patient differences relating to severity of illness and the second set addresses risk of mortality. In APR-DRGs, severity of illness is defined as the extent of organ system loss of function or physiologic decompensation, while risk of mortality is the likelihood of dying. Since severity of illness and risk of mortality are distinct patient attributes, separate subgroups are assigned to a patient for severity of illness and risk of mortality. Thus, in the APR-DRG system a patient is assigned three distinct descriptors.

- The base APR-DRG (e.g., APR-DRG 127 - Congestive Heart Failure)
- The severity of illness subgroup
- The risk of mortality subgroup

The four severity of illness subgroups and the four risk of mortality subgroups represent minor, moderate, major or extreme severity of illness or risk of mortality. The assignment of a patient to one of these four subgroups takes into consideration not only the specific secondary diagnoses but also the interaction between secondary diagnoses, age, principal diagnosis, and the presence of certain non-OR procedures. The assignment of a patient to a subgroup for each APR-DRG is performed separately for severity of illness and risk of mortality.

The creation of the four levels created the opportunity for a complete revaluation of the diagnoses categorized as a complication or comorbidity. With the expansion to four subclasses, many non-CC diagnoses could be assigned to at least a level of moderate whereas previously there was not sufficient justification to include them with the heterogeneous set of diagnoses previously categorized as a CC. As a result of this reevaluation there were 1693 diagnoses that are considered a non-CC in the AP-DRGs and Medicare DRGs that were assigned to the moderate, major or extreme level in the APR-DRGs. Conversely, 418 secondary diagnoses that are considered a CC were moved to the non-CC (minor) level in the APR-DRGs.

In the APR-DRGs, the assessment of the severity of illness or risk of mortality of a patient is specific to the APR-DRG to which a patient is assigned. In other words, the determination of the severity of illness and risk of mortality is disease specific. The significance attributed to a secondary diagnosis is dependent on the underlying problem. Thus, in APR-DRGs the severity or risk of mortality level of a secondary diagnosis is assigned separately for each base APR-DRG.

The most important component of determining the final patient subgroup is the recognition of the impact of interactions among secondary diagnoses. In APR-DRGs, high severity of illness and risk of mortality are primarily determined by the interaction of multiple diseases. Patients with multiple diseases involving multiple organ systems constitute the patients with a more extensive disease process who are difficult to treat and who have poor outcomes. In APR-DRGs specific combinations of secondary diagnoses interact causing the severity or risk of mortality subgroup to be increased.

The combination of the base APR-DRG and the final patient severity of illness and risk of mortality subgroup constitute the complete APR-DRG description of the patient. For applications such as evaluating resource use or establishing patient care guidelines, the APR-DRGs in conjunction with the severity of illness subgroup is used. For evaluating patient mortality the APR-DRG in conjunction with the risk of mortality subgroup is used.

The subdivision of the 382 base APR-DRGs into the four severity subgroups, combined with the two error APR-DRGs (i.e., 469, 470) results in 1530 APR-DRGs. The APR-DRGs were a joint development of 3M HIS and NACHRI. The APR-DRGs encompass all the DRG modifications developed for the original PM-DRGs plus all subsequent NACHRI research. APR-DRGs are updated every two years. Version 12.0 of the APR-DRGs was released in September of 1995 and was used in this analysis (Averill, Goldfield, et. al 1995).

COMPARISON OF THE STRUCTURE OF THE DRG SYSTEMS

Table 1 compares the major features of the alternative DRG systems. The five DRG systems differ substantially both with respect to the base DRGs and the subgroups within each base DRG. Base DRG categories can be formed for each DRG system by merging together all CC and age splits. The base DRG categories can not be formed for AP-DRGs because the major complication and comorbidity groups in the AP-DRGs were created by combining together patients from multiple base DRGs. Thus, in AP-DRGs it was not possible to unambiguously separate out the base DRGs from the secondary diagnosis subgroups. The number of base DRG categories is not substantially different across the DRG systems. The Medicare DRGs, SDRGs and RDRGs are focused primarily on the Medicare population while the AP-DRGs and APR-DRGs contain extensive modifications for the pediatric and adult populations. The categorization of a secondary diagnosis as a CC or non-CC was originally developed in 1980 when the ICD-9-CM Version of the DRGs was created. The Medicare DRGs and RDRGs have not reevaluated the CC list, while to varying degrees, the AP-DRGs, SDRGs and APR-DRGs have reevaluated the CC list. The APR-DRGs are the only DRG system which incorporate the impact of multiple CCs. The number of CC subgroups within a DRG varies

	Medicare DRGs Version 12.0	SDRGs	RDRGs Version 10.0	AP-DRGs Version 12.0	APR-DRGs Version 12.0
Number of Base DRG Categories	338	316	367	NA	384
Number of DRGs	492	652	1170	641	1530
Multiple Trauma MDC	Limited	Limited	Limited	Complete	Complete
HIV Infection MDC	Limited	Limited	Limited	Complete	Complete
Newborn Birthweight Used	No	No	Limited	Complete	Complete
NACHRI Pediatric Changes	No	No	No	Limited	Complete
Major (Extreme) CCs	No	Yes	Yes	Yes	Yes
Death Used In Definition	Yes	Yes	Yes	Yes	No
LOS used in Definition	No	No	Yes	Newborn Only	No
CC List Reevaluated	No	Substantial	No	Limited	Complete
Multiple CCs Recognized	No	No	No	No	Yes
Number of CC Subgroups	2	3	3 Med, 4 Surg	3	4
CC Subgroup Structure	Variable	Variable	Relatively Uniform	Variable	Uniform
Risk of Mortality subgroup	No	No	No	No	Yes
Base DRGs Used	–	Medicare	Medicare	Medicare	AP-DRGs except neonates

Table 1: Comparison of the Structure of the DRG Systems

by DRG for the Medicare DRGs, SDRGs and AP-DRGs. The CC subgroup structure for the RDRGs is relatively uniform since with the exception of a few DRGs, such as the early death DRGs which have no subgroups, there are always three subgroups for medical DRGs and four subgroups for the surgical DRGs. The CC subgroup structure in the APR-DRGs is uniform across the DRGs with all DRGs having four subgroups. A uniform CC subgroup structure facilitates the communication of comparative results to medical staffs. Death is used as a variable for assigning the DRG in all DRG systems except APR-DRGs. The use of death in the definition of the DRG makes the DRG system meaningless for mortality analysis. APR-DRGs are the only DRG system with a separate risk of mortality component.

DATABASE

The database for the study was a calendar year 1993 hospital medical record abstract discharge data set that included 675 acute general hospitals and 40 freestanding acute children's hospitals (HCIA, 1995). The national hospital sample data set used for this study included all the patients from the 675 general hospitals and a 20 percent random sample of patients from the 40 children's hospitals.

The 675 general hospitals were generally representative of acute general hospitals throughout the U.S. with respect to bed size, teaching status and urban/rural status, though there was a slight underrepresentation of rural hospitals. The 675 general hospitals included a generally representative number of hospitals from the

four Census Bureau regions of the country, with the exception of the Northeast, which was under-represented. It should be noted that this sample did not include any psychiatric hospitals or rehabilitation hospitals, though some hospitals did have psychiatric units or rehabilitation units. The 40 freestanding children's hospitals included nearly all hospitals identifying themselves as freestanding acute children's hospitals in 1993 (40 of 45).

The charge data was adjusted for area wage differences by applying the HCFA area wage index to the "labor" portion of the charges. HCFA assumes that 71 percent of charges are associated with labor. The wage index adjusted charge data was converted to an estimate of cost using a hospital-wide ratio of costs to charges (RCCs) from Medicare cost report files. This was done using the Medicare cost report values for allowable operating and capital costs, but excluding graduate medical education costs. For children's hospitals, since Medicare cost report data does not routinely exist, similar numbers were obtained from the children's hospitals.

A series of edits were applied to remove patient discharge records from the data that contained incomplete or inaccurate information. The edits eliminate patients with incalculable, invalid or unreasonable age, length of stay or charge values. For example, patients with length of stays greater than two years and patients with zero total charges were eliminated from the data. In addition, patients who were transferred to another acute care hospital were eliminated. These patients were eliminated since they represent an incomplete hospital stay. The original database contained 4,987,773 patients. The data quality edits eliminated 404,934 patients primarily as the result of missing charge information. The exclusion of the transfer patients eliminated an additional 97,745 patients. To insure a correct proportional representation of children's hospitals, only 20 percent of the patients in the children's hospitals were retained, which reduced the final number of patients in the analysis database to 4,203,646. The analysis also excluded the error DRGs 469 and 470 (in the SDRGs the error DRGs are 651 and 652). The number of patients in the error DRGs differed slightly across the five DRG systems.

UNTRIMMED R^2 RESULTS

The most common statistical measure used to compare patient classification systems is reduction of variance (R^2) which measures the proportion of variation that is explained by a DRG system. R^2 provides a summary measure of the extent to which a DRG system is able to predict the value of a resource use or outcome variable based on the characteristics of individual patients. For a categorical variable such as DRG, R^2 is computed as

$$\frac{\sum_i (y_i - A)^2 - \sum_i (y_i - A_g)^2}{\sum_i (y_i - A)^2}$$

where y_i is the value of the variable (i.e., cost or LOS) for the *ith* patient, A is the average value of the variable in the database and A_g is the average value of the variable in DRG *g*. The square of the difference between the actual value (i.e., y_i) and the predicted value (i.e., A or A_g) is a measure of the variation in the data. The term

$$\sum_i (y_i - A)^2$$

is the amount of variation before subdividing the data into DRGs and the term is the

$$\sum_i (y_i - A_g)^2$$

amount of variation after subdividing the data into DRGs. The difference between these two terms is the reduction in variation resulting from the subdivision of the data into DRGs. R^2 is the ratio of the reduction in variation to the amount of variation before subdividing into DRGs. R^2 ranges between zero and one and measures the fraction of variation explained by the DRGs. Thus, an R^2 of 0.415 would mean that subdividing the data into DRGs reduces the amount of variation in the data by 41.5 percent.

For each of the five DRG systems an R^2 for cost and LOS was computed based on untrimmed data. The R^2 was computed separately for medical patients, surgical patients and for medical and surgical patients combined. Table 2 summarizes the R^2 results.

Using cost for all patients, the percent increase in R^2 relative to the Medicare DRGs is 8.1, 13.5, 15.0 and 30.3 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The percent increase in R^2 for cost is most dramatic for medical patients at 10.6, 40.4, 50.7 and 68.9 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The APR-DRGs always perform better in terms of R^2 than the other DRG systems. For cost, the R^2 for surgical and medical patients combined is sometimes higher than either the surgical or medical patients separately. Since the cost for surgical patients tends to be much higher than medical patients, the variation before subdividing into DRGs for surgical and medical patients combined is high. Since the DRGs assign surgical and medical patients to separate groups, the reduction in variation (R^2) for surgical and medical patients combined can be higher than either the reduction in variation for medical patients or surgical patients separately.

Any classification system that subdivides data into mutually exclusive groups of patients will exhibit a higher R^2 as the number of groups increases. An R^2 of 1.0 is achieved if the number of groups equals the number of patients. Since there is a substantial difference in the number of groups in the various DRG systems, it is possible that some of the R^2 differences are due to the difference in the number of groups and not due to differences in the performance of the DRG system. The R^2 that would be achieved by randomly splitting N observations into K groups is given by the formula (Feldman, 1992):

$$\frac{K-1}{N-1}$$

Since there were 4,203,646 patients in the analysis database, the expected R^2 from randomly splitting the data into subgroups is 0.00012, 0.00016, .00028, .00015 and .00036 for the Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. Thus, the R^2 associated with the number of groups within each DRG system is negligible, and virtually none of the R^2 difference across the DRG systems is due to the difference in the number of groups in the different DRG systems.

As a means of comparing the DRG systems, the DRG structure can be viewed as four distinct components.

- MDCs
- MDCs with medical/surgical distinction
- Base DRG categories based on principal diagnosis, procedure and discharge status
- Subgroups within DRGs based primarily on secondary diagnoses and age

The MDCs with medical and surgical

Variable	Patients	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs
LOS	All	0.3126	0.3439	0.3702	0.3692	0.4213
LOS	Surgical	0.3254	0.3751	0.3757	0.3665	0.4507
LOS	Medical	0.2878	0.3046	0.3535	0.3583	0.3820
Cost	All	0.4076	0.4407	0.4627	0.4689	0.5309
Cost	Surgical	0.4111	0.4506	0.4560	0.4503	0.5218
Cost	Medical	0.2485	0.2748	0.3488	0.3744	0.4196

Table 2: Untrimmed R^2 for cost and length of stay

patients separated into distinct groups have virtually the same definition for all five systems. The base DRG categories are very similar across the DRG systems with the exception of the AP-DRGs and the APR-DRGs which contain substantial modifications to the base DRG categories. The most substantial difference in the five systems is in the definition of the subgroups within each base DRG category. The subgroups within each base DRG category are primarily based on secondary diagnoses and age, and essentially provide a measure of the relative severity of patients within each DRG category. The separation of the DRG structure into four components provides a means of identifying the source of R^2 differences across the different DRG systems. Table 3 contains the R^2 for cost across all patients for the five DRG systems for each component of the DRG structure.

As expected, the R^2 for the MDCs and the MDCs with the medical/surgical distinction is very similar across the five DRG systems and is approximately 0.11 and 0.19, respectively. At the base DRG category level, the R^2 is approximately 0.40 for all the DRG systems with the exception of APR-DRGs which has an R^2 that is about 10 percent higher than the other DRG systems. The percent increase in R^2 from the addition of the subgroups to the base DRG categories differs substantially across the DRG systems ranging from 2.69 percent for the Medicare DRGs to 21.18 percent for the APR-DRGs. Thus, it is the secondary diagnosis subgroups that account for most of the difference in R^2 across the DRG systems

R^2 FOR COST BY MDC

The untrimmed R^2 within each MDC for cost and the percent difference in R^2 relative to the Medicare DRGs is contained in

	MDC	MDC Plus Medical, Surgical	Base DRG Categories	Complete System
Medicare DRGs	0.1134	0.1852	0.3969	0.4076
SDRGs	0.1125	0.1844	0.3960	0.4407
RDRGs	0.1100	0.1823	0.4024	0.4627
AP-DRGs	0.1090	0.1916	NA	0.4689
APR-DRGs	0.1078	0.1987	0.4381	0.5309

Table 3: Untrimmed R^2 for cost for all patients for each component of the DRG structure

MDC	MDC Description	Medicare	SDRG		RDRG		AP-DRG		APR-DRG	
		R^2	R^2	% Diff	R^2	% Diff	R^2	% Diff	R^2	% Diff
1	Nervous	0.1853	0.2830	52.73	0.3006	62.22	0.2849	53.75	0.3761	102.97
2	Eye	0.0722	0.0748	3.60	0.1519	110.39	0.1303	80.47	0.2278	215.51
3	Ear, Nose, Mouth and Throat	0.1190	0.1518	27.56	0.1826	53.45	0.1998	67.90	0.2977	150.17
4	Respiratory	0.2533	0.2905	14.69	0.3142	24.04	0.2921	15.32	0.3635	43.51
5	Circulatory	0.4734	0.5315	12.27	0.5353	13.08	0.5353	13.08	0.5710	20.62
6	Digestive	0.2829	0.3535	24.96	0.3689	30.40	0.3621	28.00	0.4370	54.47
7	Hepatobiliary and Pancreas	0.1903	0.2815	47.92	0.2846	49.55	0.2627	38.05	0.3359	76.51
8	Musculoskeletal and Connective Tissue	0.2841	0.3425	20.56	0.3588	26.29	0.3564	25.45	0.4319	52.02
9	Skin, Subcutaneous Tissue and Breast	0.1729	0.2137	23.60	0.2225	28.69	0.2091	20.94	0.2821	63.16
10	Endocrine, Nutritional and Metabolic	0.1295	0.1938	49.65	0.2171	67.64	0.2088	61.24	0.3184	145.87
11	Kidney and Urinary	0.3730	0.4315	15.68	0.4391	17.72	0.3982	6.76	0.5025	34.72
12	Male Reproductive	0.2757	0.2865	3.92	0.3578	29.78	0.3103	12.55	0.3964	43.78
13	Female Reproductive	0.1645	0.2403	46.08	0.2884	75.32	0.1928	17.20	0.3424	108.15
14	Pregnancy, Childbirth and the Puerperium	0.2306	0.2150	-6.76	0.2715	17.74	0.2496	8.24	0.3110	34.72
15	Newborns	0.2917	0.2894	-0.79	0.4234	45.15	0.5070	73.81	0.6274	115.08
16	Blood and Immunological	0.0574	0.1065	85.54	0.1591	177.18	0.1368	138.33	0.2228	288.15
17	Myeloproliferative and Poorly Diff Neop	0.1810	0.2635	45.58	0.3157	74.42	0.2884	59.34	0.3543	95.75
18	Infectious and Parasitic	0.1577	0.2376	50.67	0.2624	66.39	0.2308	46.35	0.3108	97.08
19	Mental	0.0417	0.0506	21.34	0.0511	22.54	0.0427	2.40	0.1176	182.01
20	Alcohol and Drug	0.1593	0.1741	9.29	0.1158	-27.31	0.0617	-61.27	0.1549	-2.76
21	Injuries, Poisonings and Toxic Effects	0.1134	0.1993	75.75	0.1959	72.75	0.2063	81.92	0.3026	166.84
22	Burns	0.3375	0.3618	7.20	0.4209	24.71	0.3376	0.03	0.4627	37.10
23	Factors Influencing Health Status	0.1266	0.1602	26.54	0.1979	56.32	0.1210	-4.42	0.2414	90.68
24	Human Immunodeficiency Virus Infections	0.0650	0.1114	71.38	0.1603	146.62	0.2496	284.00	0.2998	361.23
25	Multiple Significant Trauma	0.1096	0.1096	-0.09	0.2032	85.23	0.2109	92.25	0.2385	117.41
26	Other (Pre MDC)	0.4070	0.4151	1.99	0.4162	2.26	0.4243	4.25	0.4721	16.49

Table 4: Untrimmed R^2 for cost across all patients by MDC

Table 4. The R^2 varies not only across the DRG systems but also across the MDCs. There is a systematic variation in R^2 across the MDCs reflecting that cost is more predictable in some MDCs such as the circulatory system and less predictable in other MDCs such as mental health. This systematic variation across MDCs reflects both the extent to which there is a relatively standard treatment for the diseases in the MDC and the need for information not currently collected such as mental health status.

The R^2 results across the individual MDCs are consistent with overall R^2 results. The APR-DRGs have the highest R^2 for all MDCs except alcohol and drugs. The magnitude of the difference in R^2 is quite substantial for many MDCs. For example, the percent increase in R^2 relative to the Medicare DRGs for MDC 24 HIV infections is 71.38, 146.62, 284.00 and 361.23 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. By weighting the percent difference in R^2 relative to the Medicare DRGs for each MDC by the number of patients in each MDC a weighted percent increase in R^2 can be computed. Table 5 contains the weighted percent difference in R^2 relative to the Medicare DRGs. The weighted percent difference in R^2 relative to the Medicare DRGs provides a composite measure of the magnitude of the difference in R^2 across MDCs. The weighted percent difference in R^2 relative to the HCFA DRGs is quite substantial. For APR-DRGs the weighted percent difference in R^2 is 163.27, 126.31 and 70.62 percent for medical patients, surgical patients and all patients combined, respectively. The weighted percent difference in R^2 across the MDCs is a better measure of the relative performance of the DRG systems than the overall R^2 since it removes the contribution of the MDCs and medical/surgical distinction which is comparable across all the DRG systems. The weighted percent increase in R^2 is lower for all patients combined than for medical and surgical patients separately because the medical/surgical distinction increases the R^2 of Medicare DRGs, lowering the percent increase in R^2 of the other DRG systems relative to the Medicare DRGs.

	SDRGs	RDRGs	AP-DRGs	APR-DRGs
All	18.78	34.26	30.45	70.62
Surgical	38.36	67.10	56.78	126.31
Medical	49.91	86.93	73.59	163.27

Table 5: Untrimmed weighted percent difference in R^2 for cost across MDCs relative to the Medicare DRGs

DATA TRIMMING

The R^2 results have been presented using untrimmed data. Data trimming excludes from the analysis patients who have extreme values of the variable being analyzed (i.e., LOS or cost). The extent and method of data trimming can have a substantial impact on R^2 and can lead to biased results. There are numerous methods for data trimming. The more common methods for establishing high and low trim points include:

- A specified number of standard deviations above and below the mean
- A specified number of interquartile ranges above the 75th percentile and below the 25th percentile
- A specified multiple of the geometric mean for the high trim points and a fraction of the geometric mean for the low trim points
- A specified percentage of the patients with the highest values of the variable

being analyzed and with the lowest values of the variable being analyzed

The first three methods of trimming would result in a different number of patients being trimmed in each of the DRG systems. The patients being trimmed are the patients who deviate the most from the value predicted by the classification system. Thus, if the R^2 of two classification systems is being compared and one of the classification systems has more patients trimmed, the R^2 results will be biased in favor of the classification system with the most patients trimmed. As an example of the magnitude of the difference in the number of patients trimmed and amount of cost trimmed across the different DRG systems a high trim point equal to four times the geometric mean of cost was used. The resulting percent of patients and cost trimmed is contained in Table 6.

As shown in Table 6, across the DRG systems the number of patients trimmed varied by more than 60 percent and the amount of cost trimmed varied by nearly 80 percent. In this example, the increase in R^2 due to trimming would be greater for the Medicare DRG than for the other DRG systems, especially the APR-DRGs. The computation of the untrimmed R^2 was based on an identical patient population across the DRG systems. Ideally, any comparison of the R^2 for trimmed data across DRG systems would also be based on an identical patient population for each system. Unfortunately, there is no trimming method that would result in an identical patient population across the DRG systems that would not be biased toward one of the systems. For example, a trimmed patient population defined by excluding patients who were trimmed in *any one* of the DRG systems or, alternatively, a trimmed patient population defined based on only those patients who were trimmed in *all* the DRG systems would create distortions within each system by not trimming patients who should be trimmed or unnecessarily excluding patients. Thus, there is no method of trimming that would produce an identical trimmed patient population across all DRG systems without creating distortions in the R^2 results.

DRG System	Percent of Patients Trimmed	Percent of Cost trimmed
Medicare DRGs	2.39	13.45
SDRGs	2.06	11.33
RDRGs	1.95	10.42
AP-DRGs	1.94	10.36
APR-DRGs	1.47	7.57

Table 6: Percent of patients and cost trimmed using a high trim of nine times the geometric mean

Since it is common for trimmed R^2 results to be reported for a DRG system, the data was trimmed by excluding within each DRG the one percent of patients with the highest values of the variable being analyzed and one half of one percent of patients with the lowest values. Thus, a total of one-and-one-half percent of the patients from each DRG were trimmed. The trimming was done independently for each DRG system. Thus, although approximately the same number of patients were trimmed from each DRG system, the patients who were trimmed differed across the DRG systems. Patients who are trimmed are referred to as outliers and patients who are not trimmed are referred to as inliers.

The patients identified as outliers differed substantially across the DRG systems. For example, the percent of high outliers for Medicare DRGs, SDRGs, RDRGs and AP-DRGs that are inliers in APR-DRGs is 54.50, 50.49, 51.90 and 50.44 percent, respectively. Thus, approxi-

mately half of the high outliers in each of these DRG systems are inliers for APR-DRGs. In order to illustrate the differences in which patients are identified as outliers across the DRG systems, Table 7 contains the high outlier patients in each DRG system that are inliers in the APR-DRG system and shows the distribution of these patients by the APR-DRG severity subclasses.

Table 7 shows that the majority of these patients are severity level 3 and 4 in APR-DRGs. Of the high outliers in each system which are inliers in APR-DRGs, 77.32, 71.87, 61.02 and 63.40 percent are severity level 3 and 4 in the APR-DRGs for the Medicare DRGs, SDRGs, RDRGs and AP-DRGs, respectively. Thus, each of the other DRG systems tend to exclude as high outliers, a disproportionate number of the high level severity patients which are inliers in APR-DRGs.

DRG System	Percent by APR-DRG Severity Subgroup			
	1	2	3	4
Medicare DRGs	8.09	14.59	24.83	52.49
SDRGs	8.48	19.64	35.55	36.32
RDRGs	9.46	29.52	29.33	31.69
AP-DRGs	8.12	28.48	37.08	26.32

Table 7: Distribution of high outliers in each DRG system that are inliers for APR-DRGs

There is greater consistency across the DRG systems in the patients who are identified as low outliers. The percent of low outliers for Medicare DRGs, SDRGs, RDRGs, and AP-DRGs that are inliers in APR-DRGs is 24.29, 23.98, 26.99 and 23.36 percent, respectively. Thus, approximately a quarter of the low outliers in each system are inliers for APR-DRGs. Table 8 contains the low outlier patients in each DRG system that are inliers in APR-DRGs and shows the distribution of these patients by the APR-DRG severity subclasses.

Table 8 shows that the majority of these are severity level 1 in APR-DRGs. Thus, each of the other DRG systems tend to exclude as low outliers a substantial number of severity level one patients which are inliers in APR-DRGs.

DRG System	Percent by APR-DRG Subgroup			
	1	2	3	4
Medicare DRGs	76.24	21.86	1.72	0.18
SDRGs	67.56	27.04	4.98	0.42
RDRGs	53.31	38.36	7.57	0.76
AP-DRGs	63.83	30.23	5.61	0.32

Table 8: Distribution of low outliers in each DRG system that are inliers for APR-DRGs

Not only are the patients defined as outliers different, but the cost of the outlier patients are also substantially different. To illustrate the difference in cost, the difference between the average inlier cost and average outlier cost for each DRG weighted by the number of patients in each DRG was computed for each DRG system. This difference is $32,395, $28,077, $27,332, $26,742 and $22,097 for the Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. Thus, there is a substantial variation in the difference between average inlier and average outlier cost across the DRG systems. For example, the weighted difference between average inlier and outlier cost is 46.6 percent higher for Medicare DRGs than for APR-DRGs. In essence, more of the unexplained variability is being excluded by trimming in the DRG systems with a greater difference between inlier and outlier average cost. The more the outliers differ from the average of the DRG to which the outlier is assigned, the greater will be the improvement in R^2. In other words, the poorer the classification system, the greater the improvement in R^2

from trimming. A system that classifies patients into highly homogeneous groups, such as the APR-DRGs, would experience relatively little increase in R^2 as a result of data trimming.

TRIMMED R^2 RESULTS

Table 9 contains the R^2 results using trimmed data. In general, the R^2 results for trimmed data are higher than for untrimmed data and the difference in R^2 across the DRG systems is smaller. For cost for all patients, the percent increase in R^2 relative to the Medicare DRGs is 3.5, 8.3, 8.7 and 16.7 percent for the SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. These are roughly half of the percent increase in R^2 relative to the Medicare DRGs observed for untrimmed data.

As expected, the trimming caused a disproportionate increase in R^2 for the DRG systems that had the lower R^2 for untrimmed data. For example, for cost across all patients, trimming resulted in a 26.4 percent increase in R^2 for HCFA DRGs (40.76 versus 51.51) and a 13.2 percent increase in R^2 for APR-DRGs (53.09 versus 60.09). As discussed previously, the reason for the greater increase in R^2 due to trimming for the DRG systems with lower untrimmed R^2 results is that the trimming excludes more of the unexplained variability in the DRG systems with lower untrimmed R^2 values. Table 10 contains the increase in R^2 due to trimming and the difference between the average inlier cost and average outlier cost for each DRG weighted by the number of patients in each DRG.

As expected, there is a direct relationship between the magnitude of the increase in R^2 and the weighted difference between the average inlier and outlier cost. Thus, even though the number of patients trimmed was approximately the same across the DRG systems, the difference in the population of patients defined as outliers and the difference in the average cost of the outliers had a substantial impact on the trimmed R^2 results. The trimmed R^2 measures both the performance of the classification system as well as the characteristics of the patients trimmed. The net impact of the trimming is that the R^2 results are artificially compressed across the DRG systems.

Table 11 contains the trimmed R^2 for cost for all patients for each component of the DRG structure. Across all DRG systems the trimmed R^2 is higher than the untrimmed R^2 for each component of the DRG systems. At the MDC level the trimmed R^2 differs across DRG systems even though the definition of the MDCs is very similar across DRG systems. The reason for the difference in R^2 is that the patients who are trimmed are different for each DRG system.

Table 13 contains the trimmed R^2 by MDC. The trimmed R^2 for each MDC is

Variable	Patients	HCFA DRG	SDRG	RDRG	AP-DRG	APR-DRG
LOS	All	0.3797	0.4070	0.4385	0.4358	0.4787
LOS	Surgical	0.3935	0.4363	0.4370	0.4289	0.5033
LOS	Medical	0.3523	0.3667	0.4259	0.4273	0.4419
Cost	All	0.5151	0.5331	0.5577	0.5600	0.6009
Cost	Surgical	0.5132	0.5342	0.5396	0.5372	0.5883
Cost	Medical	0.3432	0.3646	0.4583	0.4709	0.4981

Table 9: Trimmed R^2 for cost and length of stay

DRG System	Increase in R^2 due to Trimming	Weighted Difference in Average Inlier and Outlier Cost
Medicare DRGs	26.37%	$32,395
SDRGs	20.97%	$28,077
RDRGs	20.53%	$27,332
AP-DRGs	19.43%	$26,742
APR-DRGs	13.19%	$22,097

Table 10: Relationship between R^2 increase due to trimming and difference between inlier and outlier cost

higher than the untrimmed R^2 and the difference in R^2 across the DRG systems is smaller for trimmed data than for untrimmed data. Table 12 contains the trimmed weighted percent difference in R^2 for cost across MDCs relative to the Medicare DRGs. While the trimmed R^2 results are lower than the untrimmed results (see Table 5), there are still substantial differences across DRG systems in the trimmed weighted percent difference in R^2 for cost across MDCs relative to Medicare DRGs.

	MDC	MDC Plus Medical Surgical	Base DRG Categories	Complete System
Medicare DRGs	0.1422	0.2329	0.5024	0.5151
SDRGs	0.1361	0.2243	0.4811	0.5331
RDRGs	0.1314	0.2195	0.4862	0.5577
AP-DRGs	0.1288	0.2273	NA	0.5600
APR-DRGs	0.1203	0.2239	0.4968	0.6009

Table 11: Trimmed R^2 for cost for all patients for each component of the DRG structure

HOSPITAL LEVEL ANALYSIS FOR COST

The analysis has focused on the ability of the DRG systems to predict cost at the individual patient level. The DRGs can also be used to predict cost at the hospital level. In the previous patient level analysis the average cost of the DRG to which the patient is assigned was used to predict each individual patient's actual cost. In the hospital level analysis a casemix index based on each DRG system was computed and used to predict the actual total cost (i.e., budget) of the hospital. It is possible that one DRG system might be superior in predicting cost at the patient level but *not* be superior in predicting cost at the hospital. This would occur if the kinds of patients for which the DRG system was yielding a superior cost prediction were randomly distributed across hospitals. Conversely, if the kinds of patients for which the DRG system was yielding a superior cost production are concentrated in certain types of hospitals, there would be superior cost prediction at the hospital level.

	SDRGs	RDRGs	AP-DRGs	APR-DRGs
All	11.19	20.67	18.54	41.15
Surgical	24.15	39.15	38.24	74.22
Medical	36.66	63.80	50.69	107.61

Table 12: Trimmed weighted percent difference across MDCs in R^2 for cost relative to Medicare DRGs

Using untrimmed data a regression was performed with total hospital cost as the dependent variable and the casemix index of the hospital as the only independent variable. In addition, a multiple regression was performed using the following hospital characteristics:

Casemix Index	Urban Hospital
Percent Medicaid	Bed size less than 100
Percent Medicare	Bed size between 100 and 199
Teaching Hospital	Bed size between 200 and 299
Childrens Hospital	Bed size between 300 and 449
	Bed size 450 and greater

With the exception of the casemix index, percent Medicaid and percent Medicare

MDC	MDC Description	Medicare	SDRG		RDRG		AP-DRG		APR-DRG	
		R^2	R^2	% Diff	R^2	% Diff	R^2	% Diff	R^2	% Diff
1	Nervous	0.2903	0.3877	33.55	0.3880	33.65	0.3898	34.27	0.4512	55.43
2	Eye	0.1197	0.1272	6.27	0.1923	60.65	0.1629	36.09	0.2696	125.23
3	Ear, Nose, Mouth and Throat	0.2659	0.2914	9.59	0.2994	12.60	0.3411	28.28	0.4302	61.79
4	Respiratory	0.3625	0.3967	9.43	0.4243	17.05	0.3925	8.28	0.4562	25.85
5	Circulatory	0.6007	0.6336	5.48	0.6318	5.18	0.6265	4.29	0.6490	8.04
6	Digestive	0.4025	0.4577	13.71	0.4710	17.02	0.4588	13.99	0.5290	31.43
7	Hepatobiliary and Pancreas	0.2859	0.3690	29.07	0.3695	29.24	0.3522	23.19	0.4243	48.41
8	Musculoskeletal and Connective Tissue	0.4069	0.4380	7.64	0.4403	8.21	0.4666	14.67	0.5089	25.07
9	Skin, Subcutaneous Tissue and Breast	0.2729	0.3109	13.92	0.3021	10.70	0.3044	11.54	0.3634	33.16
10	Endocrine, Nutritional and Metabolic	0.2218	0.2826	27.41	0.2742	23.62	0.2992	34.90	0.3988	79.80
11	Kidney and Urinary	0.5030	0.5287	5.11	0.5326	5.88	0.5081	1.01	0.5797	15.25
12	Male Reproductive	0.4347	0.4310	-0.85	0.4347	0.00	0.4277	-1.61	0.4435	2.02
13	Female Reproductive	0.2520	0.3221	27.82	0.3460	37.30	0.2624	4.13	0.3881	54.01
14	Pregnancy, Childbirth and the Puerperium	0.3881	0.3809	-1.86	0.3948	1.73	0.3983	2.63	0.4262	9.82
15	Newborns	0.3569	0.3557	-0.34	0.5516	54.55	0.5726	60.44	0.6549	83.50
16	Blood and Immunological	0.1218	0.1809	48.52	0.2415	98.28	0.2157	77.09	0.2844	133.50
17	Myeloprofroliferative and Poorly Diff Neop	0.2461	0.3420	38.97	0.3644	48.07	0.3587	45.75	0.4106	66.84
18	Infectious and Parasitic	0.2343	0.3200	36.58	0.3436	46.65	0.3056	30.43	0.3834	63.64
19	Mental	0.0598	0.0733	22.58	0.0770	28.76	0.0617	3.18	0.1613	169.73
20	Alcohol and Drug	0.2530	0.2421	-4.31	0.2155	-14.82	0.0868	-65.69	0.1792	-29.17
21	Injuries, Poisonings and Toxic Effects	0.2114	0.3096	46.45	0.2942	39.17	0.3118	47.49	0.3770	78.33
22	Burns	0.3845	0.3988	3.72	0.4523	17.63	0.3845	0.00	0.4924	28.06
23	Factors Influencing Health Status	0.1748	0.2054	17.51	0.2407	37.70	0.1696	-2.97	0.2909	66.42
24	Human Immunodeficiency Virus Infections	0.1014	0.1605	58.28	0.2079	105.03	0.3178	213.41	0.3603	255.33
25	Multiple Significant Trauma	0.1519	0.1518	-0.07	0.2459	61.88	0.2599	71.10	0.2865	88.61
26	Other (Pre MDC)	0.4788	0.4799	0.23	0.4935	3.07	0.5005	4.53	0.5405	12.89

Table 13: Trimmed R^2 for cost across all patients by MDC

which are continuous variables, all variables were dummy variables (i.e., 0 or 1) indicating whether or not the characteristic applied to the hospital (e.g., was or was not an urban hospital). Only hospitals that were major teaching hospitals (interns and residents to beds ratio greater than 0.25) were considered teaching hospitals. Table 14 contains the R^2 results for cost at the hospital level for each DRG system for the casemix index alone and for the casemix index combined with the hospital characteristics listed above. Ideally, the kinds of patients a hospital treats (i.e., its casemix) would predict its cost and the addition of hospital characteristics would not increase the R^2 at the hospital level. If the addition of hospital characteristics increases R^2 then this increase may be caused by one or more of the following:

- The patient classification system may not adequately describe certain type of patients who tend to be treated in hospitals with specific characteristics.
- Patient attributes which are not directly related to the patient's medical problems (e.g., income status) may increase the cost of care for hospitals treating a disproportionate number of these patients.
- Hospital characteristics such as size may effect the cost of care independent of the casemix of the hospital (e.g., economies and diseconomies of scale).

The hospital level R^2 for the casemix index is between 20 and 30 percent higher than the patient level R^2 for cost contained in Table 2. Relative to the Medicare DRGs, the hospital level R^2 for the casemix index is 1.27, 6.71, 14.63 and 21.15 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.

The addition of the hospital characteristics increased the R^2 for all the DRG systems. The percent increase in R^2 due to the addition of the hospital characteristics was 19.56, 20.66, 15.34, 7.98 and 4.82 for the Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRG, respectively. Thus, the addition of the characteristics of the hospital does increase R^2 but for AP-DRGs and APR-DRGs the magnitude of the increase is relatively small. This means that AP-DRGs and especially APR-DRGs explain most of the difference in cost that is observed across different types of hospitals. In a payment context, AP-DRGs and APR-DRGs greatly reduce the need for payment adjustments based on hospital characteristics such as a teaching status adjustment.

PAYMENT REDISTRIBUTION

Historically, the primary use of DRGs has been as a payment mechanism. In payment applications, DRGs are the basic unit of payment and provide the basis for allocating aggregate payments to individual hospitals. The five DRG systems result in a different allocation of aggregate payments to individual hospitals. Thus, there is a redistribution of payments across hos-

	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs
R^2 Casemix Index Only	0.5292	0.5358	0.5646	0.6065	0.6410
R^2 Casemix Index Plus Hospital Characteristics	0.6326	0.6465	0.6512	0.6549	0.6719
Percent Increase in R^2 From Hospital Characteristics	19.56	20.66	15.34	7.98	4.82

Table 14: Untrimmed R^2 for cost at hospital level across DRG systems

pitals associated with use of the alternative DRG systems. The extent to which the payment redistribution impacts certain groups of patients or hospitals is an important factor in the evaluation of alternative DRG systems. It is important not only for payment applications, but also for hospital profiling by state data commissions, planning and budgeting by hospital systems, and utilization review and pricing by individual hospitals.

In order to measure the payment redistribution impact, hospital payments under each DRG system were simulated and compared to hospital cost. DRG payment systems can be complex involving adjustments for factors such as teaching status and separate payments for outliers. The payment redistribution analysis simulated a basic DRG payment system in which payment for a patient is equal to the average cost in the DRG to which the patient is assigned without any adjustment factors or separate payments for outliers.

A basic DRG payment model was used because the purpose of the analysis was to evaluate the performance of the DRG classification systems themselves and not the additional payment adjustment polices which are typically a part of a prospective payment system. Secondarily, the use of a basic DRG payment model provides some insight as to where payment system adjustments may be most needed and whether such adjustments are less important for the more refined DRG systems.

The payment redistribution impact was measured by the ratio of payment to actual cost (RPAC) which is computed as

$$RPAC_s = \frac{\sum_g n_{g,s} A_g}{\sum_{i \in s} y_i}$$

where s is the subset of patients being evaluated (e.g., patients with age 80 or greater), y_i is the actual cost of the i^{th} patient, $n_{g,s}$ is the number of patients in DRG g in subset s and A_g is the average cost in DRG g in the entire database. The numerator of RPAC is equivalent to the total amount a hospital would be paid if the payment it received for each patient was equal to the average cost in the DRG assigned to the patient. The computation of the RPAC is budget neutral since across all patients the value of RPAC will be 1.00.

The value of RPAC provides two comparative measures. First, RPAC provides a measure of the extent to which hospital payments correspond to actual hospital cost. Values of RPAC near 1.00 mean that hospital payments closely correspond to hospital costs. Thus, DRG systems that have RPAC values close to 1.00 provide payments that are close to historical hospital cost. Second, the relative value of RPAC provides a measure of the redistribution impact of the different DRG systems. For example, suppose for a certain subset of hospitals the RPAC value for one DRG system was 1.20 and for another DRG system it was 1.10. Under the first DRG system the subset of hospital would be paid 20 percent more than cost and under the second DRG system they would be paid 10 percent more than cost. Relative to the first DRG system, payments to the subset of hospitals under the second DRG system would be 8.3 percent less ((1.1-1.2)/1.2). The primary focus of the payment redistribution analysis is to identify whether certain DRG systems produce RPAC values closer to 1.00 than other DRG systems. It is the direction and magnitude of the difference in RPAC that is critical, not whether a value of 1.00 is produced.

The degree to which hospital DRG based payments correspond to hospital costs and the payment redistribution impact of alternative DRG systems are critical issues. If certain hospitals dominate the care rendered to patients for which DRG payments are significantly less than cost, then these hospitals will financially be disadvantaged. This is especially important in a highly managed care context where health plans are actively directing inpatient admissions to selected hospitals which may appear more price competitive. Ultimately, there may be financial pressure to not offer selected services. The result will be that groups of patients might find it difficult to access needed hospital services.

The RPAC was calculated for different subsets of patients and hospitals for each of the five DRG systems. Payment redistribution was measured by computing the percent difference in RPAC relative to the Medicare DRGs. A percent difference of +20 percent means that the subset of patients or hospitals are paid 20 percent more under the specified DRG system than under the Medicare DRGs. For brevity of presentation, the tables show the RPAC for each DRG system but only show the percent difference of APR-DRGs relative to Medicare DRGs. In most instances, the largest percent difference occurs with the APR-DRGs.

Patients with discharge destination of transferred to acute hospitals were removed from the analysis database. This is because they do not have a complete hospitalization. Thus, the analysis of payment redistribution is not complicated by the issue of how to handle transfer patients.

Payment redistribution results are presented for the following subsets of patients and hospitals:

- Age
- Admission source
- Discharge destination
- Primary payor
- Hospital service line and age
- Hospital type and age

RPACs were also generated for other combinations of variables including: admission source and age, discharge destination and age, and primary payor and age. For brevity, these tables are not included, but the most salient patterns are incorporated into the presentation of findings.

Age Range

Table 15 contains the RPAC for five different patient age ranges. The pediatric age range also includes a separate breakout for children with chronic diseases and children without chronic diseases. Chronic is defined as a disease expected to last 12 months or longer for at least 75 percent of the patients. The definition of chronic was based on Version 3.0 of the NACHRI classification of congenital and chronic conditions which included physical conditions but not mental health conditions.

There are substantial differences in RPAC by patient age. For five of the six age ranges, the RPAC is closer to 1.00 for the more refined DRG systems, in particular, the APR-DRGs. For neonates, the RPAC increases from approximately 0.900 for Medicare DRGs, SDRGs and RDRGs up to approximately 1.00 for AP-DRGs and APR-DRGs. In the Medicare DRGs, SDRGs and RDRGs neonates are defined by the presence of certain newborn and perinatal diagnoses codes which results in certain very expensive neonates being assigned to non neonatal DRGs. In contrast, the AP-DRGs and APR-DRGs define neonates by age (0-28 days at admission) and therefore more appropriately classify

Patient Age Groups	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Neonates	492,558	1,066,290	0.899	0.893	0.914	0.994	1.000	+11.2%
Patients age 1 mo-17 years								
With chronic DX	115,608	888,040	0.824	0.816	0.853	0.849	0.898	+9.0%
Without chronic DX	195,871	673,475	1.105	1.076	1.124	1.067	1.036	-6.2%
Subtotal	311,479	1,561,515	0.945	0.928	0.970	0.943	0.957	+1.3%
Patients 18-64 years	1,953,661	10,813,169	1.029	1.019	1.014	1.019	0.990	-3.8%
Patients 65-79 years	956,908	8,002,197	0.982	0.988	0.987	0.976	0.996	+1.4%
Patients 80+ years	472,641	3,386,990	1.009	1.034	1.028	1.025	1.063	+5.4%

Table 15: Comparison of RPAC for DRG systems by patient age group

the neonatal patients.

The RPACs show that hospitals are substantially underpaid for children with chronic diseases and substantially overpaid for children who do not have chronic diseases. The RPACs for Medicare DRGs are 0.824 for chronically ill children and 1.105 for non chronically ill children. The RPACs move somewhat closer to 1.00 with the SDRGs, RDRGs and AP-DRGs but only by 2 percent to 3 percent. With APR-DRGs, the RPAC for chronically ill children increases 9.0 percent to 0.898 and the RPAC for non chronically ill children decreases 6.2 percent to 1.036. Thus, with the APR-DRGs the underpayment for chronically ill children is substantially decreased but a significant gap remains. This means that hospitals which specialize in the care of chronically ill children will be underpaid by all DRG systems, but will get paid substantially closer to cost with APR-DRGs.

The RPAC for adult patients age 18-64 years is a little above 1.00 for Medicare DRGs, SDRGs, RDRGs and AP-DRGs, but decreases 3.8 percent to 0.990 for APR-DRGs. This reflects a lower distribution of severity subclasses in the APR-DRGs for adults. The RPAC for elderly patients 65-79 years increases 1.4 percent from 0.982 for Medicare DRGs to 0.996 for APR-DRGs reflecting a slightly higher distribution of higher severity subclasses in the APR-DRGs for elderly patients.

Elderly patients age 80 years and older show a different pattern with RPACs having values above 1.00 which increase with the more refined DRG systems. The RPAC increases 5.4 percent from 1.009 for Medicare DRGs to 1.063 for APR-DRGs. It is likely that this increase reflects two factors. First, a higher proportion of this age population will have organ failure diagnoses placing them in high severity subclasses. Second, it is likely that a higher proportion of this age population will receive treatment aimed more at comfort than cure, including “Do not resuscitate” orders.

Overall, the more refined DRG systems decrease payment for adults and pediatric patients without chronic illnesses and increase payments for neonates, children with chronic diseases and elderly patients. In the more refined DRG systems the RPACs are closer to 1.0 for all age ranges except elderly patients age 80 years and older.

Admission Source

Table 16 contains RPACs by patient

Patient Admission Source	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Other/unknown	4,034,699	23,101,242	1.010	1.008	1.008	1.006	1.003	-0.7%
Acute hospital	119,186	1,485,889	0.849	0.874	0.876	0.899	0.930	+9.5%
SNF	33,362	243,031	0.997	1.053	1.025	1.058	1.113	+11.6%

Table 16: Comparison of RPAC for DRG systems by patient admission source

admission source. The patterns are very striking. With the more refined DRG systems the RPACs are much closer to 1.00 for patients who are transferred in from another acute hospital but are further from 1.0 for patients transferred in from a skilled nursing facility (SNF).

The RPAC for patients transferred in from another acute hospital is 0.849 for Medicare DRGs, increases to 0.874 and 0.876 for SDRGs and RDRGs to 0.899 for AP-DRGs, and to 0.930 for APR-DRGs. It is likely that these are very sick patients. It is also likely that the patients transferred from another acute hospital are concentrated in certain types of hospitals and thus, it is very important that this be reflected in the DRG system. For example, the RPAC for neonates who are transferred from another acute hospital increases 73.0 percent from 0.548 for Medicare DRGs to 0.948 for APR-DRGs.

The RPAC for patients who are transferred in from a SNF is 0.997 for Medicare DRGs, increases to 1.025 for RDRGs, to 1.053 and 1.058 for SDRGs and AP-DRGs, and to 1.113 for APR-DRGs. The majority of these patients are elderly and it is likely that the higher RPACs, especially for APR-DRGs, reflects that many of these patients have organ failure and other severe complicating or comorbid conditions but do not always receive aggressive cure oriented treatment, especially when “Do not resuscitate” orders are given.

Discharge Destination

Table 17 contains the RPACs by patient discharge destination. Again there are some rather striking patterns with the RPACs closer to 1.00 for the more refined DRG classification systems, especially the APR-DRGs. The RPAC for patients discharged to home decreases 4.7 percent from 1.085 for Medicare DRGs to 1.034 for APR-DRGs. The RPAC for patients who left against medical advice decreases 5.4 percent from 1.313 to 1.242. This suggests the APR-DRG system is classifying many of these patients in lower severity subclasses. The RPAC is still very high, however. This is not surprising given that these patients, by their choice, do not receive the full treatment that would normally be provided.

The RPAC for patients discharged to home health service, skilled nursing facility (SNF), intermediate care facility (ICF), and other institutions (e.g., rehabilitation facility) show a similar pattern to each other. The RPACs are in the range of 0.796 to 0.890 for Medicare DRGs, increase 3 to 5 percent for SDRGs, RDRGs and AP-DRGs, and increase 7 to 10 percent for APR-DRGs resulting in RPACs in the range of 0.855 to 0.957 for APR-DRGs. It is likely that these patients with ongoing treatment needs, whether in the home setting or other facility, have a higher severity and the higher RPAC for the more refined DRG systems, especially the APR-DRGs, is appropriate.

Patient Discharge Destination	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Home	3,475,196	16,797,798	1.085	1.062	1.066	1.064	1.034	-4.7%
Left against medical advice	28,888	101,002	1.313	1.300	1.290	1.283	1.242	-5.4%
Home health service	241,555	2,544,250	0.860	0.882	0.894	0.879	0.919	+6.9%
Other institution	92,643	1,074,774	0.796	0.825	0.836	0.818	0.855	+7.4%
SNF	201,238	2,189,539	0.821	0.869	0.865	0.856	0.907	+10.5%
ICF	33,802	283,399	0.890	0.923	0.918	0.914	0.957	+7.5%
Died	113,128	1,833,529	0.745	0.852	0.800	0.860	0.992	+33.2%
Unknown	797	5,871	0.969	1.060	1.038	1.059	1.162	+20.0%

Table 17: Comparison of RPACs for DRG systems by patient discharge destination

The most striking discharge destination pattern is for patient who die. The RPAC increases from 0.745 for Medicare DRGs, to 0.800 for RDRGs, to 0.852 and 0.860 for SDRGs and AP-DRGs, and to 0.992 for APR-DRGs. This primarily reflects the experience of the elderly (65+ years) who account for three-quarters of the inpatient hospital deaths. The RPACs for patients who die are very variable through the different age ranges. For example, chronically ill children for whom the RPAC is a very low 0.385 with Medicare DRGs, increases to 0.433, 0.455, and 0.487 for SDRGs, RDRGs and AP-DRGs respectively, and to 0.589 for APR-DRGs. It is important to realize that the cost of patients who die is extremely variable. There are patients who die within the first or second day of admission, those who die after many weeks of intensive treatment and terminally ill patients who receive only comfort oriented treatment.

Overall, the more refined DRG systems decrease payment for patients discharged home and patients who left against medical advice and increase payment for all other discharge destinations.

Primary Payor

Table 18 contains the RPAC by the primary payor of the patient. There are two patterns. First, the RPACs are closer to 1.00 for the more refined DRG systems, especially the APR-DRGs, for all primary payors except Other Government for which the RPACs are slightly further from 1.00. Second, the RPAC for Medicare, Medicaid and Title V (Maternal and Child Health) increases 2 to 3 percent from values that are less than 1.00. For other primary payors (except Other Government), the RPACs decrease 2 to 6 percent from values that are greater than 1.00. In other words, the Medicare DRGs tend to underpay Medicare, Medicaid and Title V patients and overpay most other payors. The more refined DRG systems, especially the APR-DRGs, provide more accurate payment levels for different primary payors.

For Medicare patients the RPAC increases 2.8 percent from 0.983 for Medicare DRGs to 1.011 for APR-DRGs. The largest increases in RPAC for Medicare patients are for chronically ill children, presumably many of whom are ESRD patients, followed by patients age 18-64, presumably many of whom are SSI Disabled, and then by patients age 80+ years, and patients age 65-79 years. RPACs for chronically ill children and 18-64 year old

Patient Third Party Payor	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Medicare	1,457,912	11,610,304	0.983	0.996	0.993	0.984	1.011	+2.8%
Medicaid	705,439	3,075,711	0.921	0.918	0.940	0.948	0.941	+2.2%
Title V	648	6,434	0.734	0.720	0.748	0.821	0.839	+14.3%
Other government	55,300	324,592	0.996	0.988	0.989	0.998	0.975	-2.1%
Workman's compensation	31,697	205,526	1.036	1.008	1.010	1.023	0.968	-6.6%
Blue Cross/insur/other/ no charge	1,722,248	8,558,782	1.046	1.033	1.028	1.035	1.008	-3.6%
Self-pay	214,003	1,048,813	1.033	1.026	1.027	1.036	1.007	-2.5%

Table 18: Comparison of the RPAC for DRG systems by primary payor

Medicare patients are very low for Medicare DRGs and increase closer to but not up to 1.00 for the more refined DRG systems.

For Medicaid patients, the RPACs increase from 0.921 for Medicare DRGs to 0.941 for the APR-DRGs. The largest increases in RPACs are for neonatal patients and chronically ill children, both of whom have very low RPACs for Medicare DRGs which remain low even for the more refined DRG systems.

The decrease in RPAC for patients with nongovernment payors was largest with the APR-DRGs. Of all the primary payors, the decrease in RPAC was largest for Workman's Compensation patients with a 6.6 percent decrease followed by Blue Cross, Commercial Insurance and No Charge patients with a 3.6 percent decrease and Self-Pay patients with a 2.5 percent decrease.

Hospital Service Line and Age

Table 19 contains the RPAC by service line and patient age group. For this analysis, five service lines were defined. Table 19 contains the results for each age group within each service line for which there was appreciable patient volume.

The methodology for defining service lines was the following hierarchical assignment logic:

- *Normal Newborn* - Patients classified as normal newborn in any of the five DRG systems
- *Mental Health/Drug & Alcohol Abuse* - Patients classified to the mental health and drug & alcohol abuse DRGs in any of the five DRG systems
- *Obstetrical* - Patients classified to an obstetric DRG in any of the five DRG systems
- *Surgical* - Patients classified to a surgical DRG in any of the five DRG systems
- *Medical* - All patients not classified into one of the prior four service lines

For normal newborns the RPAC decreases 17.5 percent RPAC from 1.231 for Medicare DRGs to 1.016 for APR-DRGs. In Medicare DRGs and SDRGs, many neonates who are really normal newborns are considered to have significant problems, and as a result, the RPAC is as high as 1.231. While these newborns are relatively inexpensive and dollar differences per case are small, total case volume is very large resulting in a significant financial impact. Approximately one-fourth of the newborns classified by the APR-DRGs as normal newborns are classified by Medicare DRGs and SDRGs as having significant problems.

	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Normal Newborn								
Normal Newborn	427,131	264,942	1.231	1.221	1.028	0.988	1.016	-17.5%
Mental Health/Drug & Alcohol Abuse (MDCs 19 & 20)								
1 mo-17 yrs w/chronic DX	1,722	13,338	0.641	0.638	0.770	0.640	0.669	+4.4%
1 mo-17 yrs w/o chronic DX	13,351	78,444	0.789	0.785	0.986	0.780	0.727	-7.9%
18-64 yrs	124,633	497,177	1.110	1.103	1.066	1.102	1.073	-3.3%
65-79 yrs	19,512	122,177	0.794	0.786	0.761	0.788	0.882	+11.1%
80+ yrs	9,643	57,813	0.868	0.865	0.842	0.862	0.961	+10.7%
Subtotal	168,861	768,949	1.000	0.994	0.987	0.993	0.992	-0.8%
Obstetrics (MDC 14)								
1 mo-17 yrs w/chronic DX	679	1,954	0.729	0.724	0.804	0.722	0.841	+15.4%
1 mo-17 yrs w/o chronic DX	23,902	47,965	0.981	0.975	0.978	0.974	0.979	-0.2%
18-64 yrs	384,210	750,279	1.015	1.008	1.008	1.008	1.004	-1.1%
Subtotal	408,791	800,198	1.013	1.006	1.005	1.006	1.002	-1.1%
Surgical								
Neonates	3,880	209,068	0.319	0.316	0.403	0.869	1.004	+214.7%
1 mo-17 yrs w/chronic DX	30,473	450,259	0.850	0.829	0.856	0.855	0.899	+5.8%
1 mo-17 yrs w/o chronic DX	39,400	239,053	1.142	1.068	1.089	1.109	1.037	-9.2%
18-64 yrs	691,629	6,004,421	1.037	1.025	1.021	1.025	0.997	-3.9%
65-79 yrs	338,344	4,505,992	0.990	1.003	0.996	0.983	1.004	+1.4%
80+ yrs	110,956	1,405,391	0.997	1.040	1.038	1.109	1.063	+6.6%
Subtotal	1,214,682	12,814,184	1.000	1.000	1.000	1.002	1.004	+0.4%
Medical								
Neonates	61,452	591,673	0.956	0.950	1.044	1.041	0.991	+3.7%
1 mo-17 yrs w/chronic DX	82,734	422,489	0.802	0.808	0.853	0.851	0.904	+12.7%
1 mo-17 yrs w/o chronic DX	119,207	307,994	1.176	1.172	1.209	1.121	1.122	-4.6%
18-64 yrs	753,189	3,561,291	1.006	1.000	0.997	1.000	0.963	-4.3%
65-79 yrs	599,050	3,374,022	0.977	0.977	0.982	0.973	0.988	+1.1%
80+ yrs	352,039	1,923,781	1.022	1.034	1.026	1.033	1.067	+4.4%
Subtotal	1,967,671	10,181,250	0.993	0.993	1.000	0.997	0.996	+0.3%

Table 19: Comparison of RPAC for DRG systems by hospital service line and patient age

Across all age ranges, the RPAC for mental health and drug and alcohol patients is often substantially different from 1.00. Costs are generally underpaid for pediatric and elderly patients and over paid for adult patients. There is one DRG system, the RDRGs, which uses age 0-17 years instead of secondary diagnoses to define mental health subclasses and this largely removes the underpayment for children except for those who also have physical chronic diseases. There appears to be distinctive mental health/drug and alcohol abuse problems that present at different ages and the more refined DRG systems have RPACs closer to 1.0 but the extent of the improvement is limited given the available diagnostic information.

For adult obstetrical patients the RPAC decreased 0.9 percent from 1.015 for Medicare DRGs to 1.004 for APR-DRGs. For pediatric obstetrical patients without chronic conditions the RPAC remains essentially the same with a value of 0.981 for Medicare DRGs and 0.979 for APR-DRGs. For pediatric obstetrical patients with chronic conditions the RPAC increases 15.4 percent from 0.729 for Medicare DRGs to 0.841 for APR-DRGs. In general, adult obstetric patients have RPACs slightly above 1.00 and pediatric obstetrical patients have RPACs below 1.00.

Although overall surgical patients have RPACs very close to 1.00 for all DRG systems, there is substantial variation by patient age. The difference is most dramatic for neonatal surgical patients with an increase in RPAC of 214.7 percent from 0.319 for Medicare DRGs to 1.004 for APR-DRGs. From the stand point of financial risk, this is very important since only a very small number of hospitals offer neonatal surgical services. The RPAC for surgical patients increases for the more refined DRG systems for pediatric patients with chronic conditions and patients over 64 and decreases for pediatric patients without chronic conditions and patients age 18-64.

Overall medical patients have RPACs very close to 1.00 for all DRG systems, but there is substantial variation by patient age. The RPAC for medical patients increases for the more refined DRG systems for neonates, pediatric patients with chronic conditions and patients over 64 and decreases for pediatric patients without chronic conditions and patients age 18-64.

Hospital Type and Age

Table 20 contains RPACs by hospital type and patient age. Hospital type is defined as follows:

- Children's hospital (number = 40)
  Based on 20 percent random sample of patients from children's hospitals
- Major teaching hospital (number = 28)
  Hospitals with a ratio of interns and residents to beds ≥ 0.25
- Other urban hospital (number = 413)
  Excludes children's and major teaching hospitals
- Other rural hospitals (number = 234)
  Excludes children's and major teaching hospitals

The refined DRG systems, and especially the APR-DRGs, produce RPACs that are much closer to 1.00 for all four hospital types. The percent change is greatest for those hospital types for which the RPAC for Medicare DRGs is furthest from 1.00.

For children's hospitals, the RPAC increases 20.4 percent from 0.691 for Medicare DRGs to 0.832 for APR-DRGs. The percent increase is 106.8 percent for neonates, nearly all of whom are transfers in or readmissions, and 10.5 percent for

Hospital Type	Total Cases	Total Costs (000s)	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs	%Diff
Children's Hospitals								
Neonates	5,751	148,548	0.440	0.440	0.485	0.769	0.910	+106.8%
1Mo-17 yrs w/chronic DX	36,266	368,130	0.715	0.706	0.728	0.739	0.790	+10.5%
1Mo-17 yrs w/o chronic DX	27,627	121,587	0.923	0.893	0.903	0.888	0.866	-6.2%
18+ yrs	389	3,100	0.771	0.707	0.669	0.763	0.672	-12.3%
Subtotal	70,033	641,365	0.691	0.680	0.705	0.774	0.832	+20.4%
Major Teaching Hospitals								
Neonates	61,906	291,737	0.623	0.611	0.785	0.871	0.892	+43.2%
1Mo-17 yrs w/chronic DX	27,203	249,267	0.819	0.814	0.850	0.853	0.920	+12.3%
1Mo-17 yrs w/o chronic DX	27,121	133,424	0.914	0.887	0.910	0.890	0.885	-3.2%
18-64 yrs	282,160	2,286,074	0.888	0.882	0.884	0.889	0.879	-1.0%
65-79 yrs	73,533	853,848	0.862	0.878	0.874	0.868	0.894	+3.7%
80+ yrs	23,275	227,647	0.840	0.872	0.862	0.865	0.904	+7.6%
Subtotal	495,198	4,041,997	0.857	0.857	0.872	0.880	0.887	+3.5%
Other Urban Hospitals								
Neonates	368,327	583,296	1.119	1.113	1.064	1.099	1.068	-4.6%
1Mo-17 yrs w/chronic DX	45,574	252,123	0.953	0.945	1.002	0.974	1.002	+5.1%
1Mo-17 yrs w/o chronic DX	113,686	350,938	1.195	1.165	1.235	1.155	1.112	-6.9%
18-64 yrs	1,450,127	7,648,775	1.055	1.045	1.040	1.044	1.012	-4.1%
65-79 yrs	754,834	6,360,267	0.984	0.991	0.990	0.977	0.999	+1.5%
80+ yrs	375,489	2,728,092	1.007	1.032	1.028	1.023	1.066	+5.9%
Subtotal	3,108,037	17,923,490	1.026	1.027	1.024	1.020	1.019	-0.7%
Other Rural Hospitals								
Neonates	56,574	42,710	1.388	1.386	1.231	1.195	1.122	-19.2%
1Mo-17 yrs w/chronic DX	6,565	18,520	1.286	1.271	1.360	1.290	1.309	+1.8%
1Mo-17 yrs w/o chronic DX	27,437	67,527	1.340	1.318	1.366	1.278	1.245	-7.1%
18-64 yrs	220,995	875,330	1.165	1.149	1.131	1.146	1.084	-7.0%
65-79 yrs	128,535	788,039	1.094	1.086	1.082	1.078	1.074	-1.8%
80+ yrs	73,873	431,184	1.112	1.128	1.112	1.122	1.130	+1.6%
Subtotals	513,979	2,223,310	1.140	1.133	1.121	1.123	1.097	-3.8%

Table 20: Comparison of RPAC for DRG systems by hospital type and patient age

pediatric patients with chronic conditions. The percent decrease is 6.2 percent for pediatric patients without chronic conditions.

For major teaching hospitals, the RPAC increases 3.5 percent from 0.857 for Medicare DRGs to 0.887 for APR-DRGs. The percentage increase in RPAC is 4.6 percent for neonates, 12.3 percent for pediatric patients with chronic conditions, 3.7 percent for patients age 65-79 and 7.6 percent for elderly patients 80 years and older. The percentage decrease in RPAC is 3.2 percent for pediatric patients without chronic conditions and 1.0 percent for adults age 18 to 64. The RPAC results for pediatric patients are fairly similar for major teaching hospitals and children's hospitals.

For other urban hospitals, the RPAC decreased 0.7 percent from 1.026 with Medicare DRGs to 1.019 with APR-DRGs. The percent decrease is 4.6 percent for neonates, 6.9 percent for pediatric patients without chronic conditions and 4.1 percent for adults age 18-64. The RPAC increases 5.1 percent for neonates, 1.5 percent for patients age 65-79 and 5.9 percent for elderly patients age 80 years and older.

For other rural hospitals, the RPAC decreases 3.8 percent from 1.140 for Medicare DRGs to 1.097 with APR-DRGs. The percent increase in RPAC is 1.8 percent for pediatric patients with chronic conditions and 1.6 percent for elderly patients 80 years and older. The percent decrease in RPAC is 19.2 percent for neonates, 7.1 percent for pediatric patients without chronic conditions, 7.0 percent for patients age 18-64 and 1.8 percent for patients age 65-79.

Overall, the more refined DRG systems and especially the APR-DRGs, generate RPACs that are substantially closer to 1.00 for most age ranges and hospital types, in particular for children's hospitals and major teaching hospitals. Thus, the more refined DRG systems redistribute payment so that the extent of under and over payment is significantly reduced. This greatly reduces the need for payment adjustments such as teaching adjustments which are common in most DRG payment systems. It also provides classifications that are far more reliable for a variety of purposes including hospital profiling, planning and budgeting, utilization review and pricing.

MORTALITY ANALYSIS

The purpose of the mortality analyses is to compare the ability of the various DRG systems to predict inpatient mortality. The DRGs, especially the APR-DRGs are increasingly being used for this purpose. Numerous states are now publishing hospital risk adjusted mortality rates in order to permit consumers to compare hospital outcomes.The Medicare DRGs, SDRGs, RDRGs and AP-DRGs systems were developed for predicting resource use and were not intended for mortality prediction. In these systems mortality is one of the variables used to define the DRGs (i.e., patients are assigned to different DRGs depending on whether they lived or died). This is not appropriate for a mortality prediction model since it would be circular logic to use mortality to predict mortality. Therefore, for the mortality analyses, the data were re-grouped eliminating all mortality distinctions, (i.e., patients are grouped into the DRG to which they would have been assigned if the patient had not died). Since APR-DRGs do not use mortality as a grouping variable, re-grouping was not necessary for APR-DRGs.

The APR-DRGs have a separate set of severity subclasses which group patients based on the risk of mortality. These APR-DRG risk of mortality subclasses were used for the mortality analyses instead of the APR-DRG severity subclasses which were used for the analyses of cost and length of stay.

Mortality in the hospital is used in this analysis. Mortality in the hospital is commonly available in hospital administrative records, but if the patient dies the day following discharge from the hospital, this would not be reflected in the hospital's records. Ideally, mortality subsequent to discharge would have been merged with the data, but this information was not available.

The literature which assesses the ability of various models to predict mortality relies on two basic statistics, R^2 and the area under the receiver operating characteristics (ROC) curve. In order to be consistent with this literature, the same two statistics were used for evaluating the ability of the different DRG systems to predict inpatient mortality. The R^2 for mortality is computed by assigning each patient a value of zero or one indicating whether they were discharged alive or dead, respectively. The predicted mortality for the patient is equal to the average value of the zero/one variable in the DRG to which the patient is assigned. The average value of the zero/one value is equivalent to the fraction of patients who died in the DRG. Based on the zero/one variable, the R^2 for mortality is computed in the same manner as the R^2 for cost or length of stay.

The area under the receiver operating characteristics (ROC) curve, is commonly used to evaluate alternative methods for predicting a zero/one outcome. The area under the ROC curve is typically used for evaluating the efficacy of a method which predicts that a given patient will or will not experience the event of interest. The basis of the ROC curve is sensitivity and specificity. In this context, sensitivity is the probability that someone classified as dead actually died, while specificity is the probability that someone classified as alive actually did not die. Sensitivity and specificity are computed as follows:

$$\text{Sensitivity} = \Sigma n_i p_i \, I(p_i \geq P)/\Sigma n_i p_i$$

$$\text{Specificity} = \Sigma n_i (1-p_i) \, I(p_i < P)/\Sigma n_i (1-p_i)$$

where n_i is the number of patients in DRG *i*, p_i is the fraction of patients who died in DRG *i* and $I(p_i)$ is an indicator that takes the value 1.0 if, in a particular DRG, the proportion of dead is at least P for sensitivity and less than P for specificity. The computation of the sensitivity and specificity assumes that all patients in a DRG died if $p_i \geq P$ and conversely that all patients in a DRG lived if $p_i < P$.

The ROC curve plots sensitivity against one minus specificity as the value of P varies. The area under the ROC curve is referred to as the c-statistic. The c-statistic measures how well a DRG system discriminates between patients who lived and those who died. A c-statistic value of 0.5 indicates no ability to discriminate while a value of 1.0 indicates perfect discrimination.

In order to understand the interpretation of the c-statistic assume that patients are separated into two groups comprising those who died and those who did not. If a patient is drawn from each group at random, then each of these patients will have an associated DRG and each DRG will have an associated mortality rate (i.e., fraction of patients who die). The c-statistic is the probability that the mortality rate in the DRG assigned to the patient who died is higher than the mortality rate in the

Statistic	Patients	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs
R^2	All	.1076	.1662	.1532	.1507	.2638
R^2	Surgical	.0773	.1467	.1247	.1258	.2431
R^2	Medical	.1149	.1707	.1601	.1567	.2676
c	All	.8926	.9230	.9189	.9252	.9481
c	Surgical	.8810	.9315	.9238	.9275	.9570
c	Medical	.8935	.9186	.9157	.9228	.9439

Table 21: R^2 and c-statistics for mortality

DRG assigned to the patient who lived. This probability would be expected to be high (i.e., near 1.0) for all DRG systems since the deaths tend to be concentrated in relatively few DRGs. Thus, the c-statistic should be relatively high for each DRG system and as a result, is unlikely to provide much differentiation across the DRG systems.

Table 21 contains the R^2 and c-statistic for mortality. Across all patients, the percent increase in R^2 relative to the Medicare DRGs is 54.5, 42.4, 40.1 and 145.2 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. As expected, the c-statistic was high for all DRG systems. Across all patients, the percent increase in the c-statistic relative to the Medicare DRGs is 3.4, 2.9, 3.7 and 6.2 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.

Table 22 contains the R^2 and c-statistic for mortality for each DRG system for each component of the DRG structure. The R^2 and c-statistic for mortality at the MDC and DRG category level are virtually the same across the DRG systems. The percent increase in R^2 from the addition of the subgroups to the DRG categories ranges from 9.02 percent for the Medicare DRGs to 141.35 percent for the APR-DRGs. The percent increase in the c-statistic from the addition of the subgroups to the DRG categories ranges from 1.40 percent for the Medicare DRGs to 7.15 percent for the APR-DRGs. Thus, it is the secondary diagnosis subgroups that account for virtually all of the difference in R^2 and c-statistic across the DRG systems.

Statistic	DRG System	MDC	MDC Plus Medical, Surgical	DRG Categories	Complete System
R^2	Medicare DRGs	0.0248	0.0263	0.0987	0.1076
R^2	SDRGs	0.0249	0.0264	0.1024	0.1662
R^2	RDRGs	0.0247	0.0262	0.1019	0.1532
R^2	AP-DRGs	0.0247	0.0264	NA	0.1507
R^2	APR-DRGs	0.0247	0.0266	0.1093	0.2638
c	Medicare DRGs	0.7635	0.7720	0.8803	0.8926
c	SDRGs	0.7639	0.7723	0.8834	0.9230
c	RDRGs	0.7635	0.7720	0.8843	0.9189
c	AP-DRGs	0.7629	0.7723	NA	0.9252
c	APR-DRGs	0.7632	0.7732	0.8848	0.9481

Table 22: R^2 and c-statistic for mortality for each component of the DRG structure

R^2 AND C-STATISTIC FOR MORTALITY BY MDC

The R^2 within each MDC for mortality and the percent difference in R^2 relative to the Medicare DRGs is contained in Table 23. The c-statistic within each MDC for mortality and the percent difference in the c-statistic relative to the Medicare DRGs is contained in Table 24.

The magnitude of the difference in R^2 for mortality across the DRG systems can be very large for some MDCs. Since R^2 for mortality for the Medicare DRGs is very low for some MDCs, the difference in R^2 can be several thousand percent. Since the c-statistic is relatively high for all DRG systems, the magnitude of the difference in the c-statistic across the DRG systems tends to be in the 5-20 percent range for most MDCs. For example, the percent increase in the c-statistic for the APR-DRGs relative to the Medicare DRGs ranges 2.75 percent for newborns to 40.25 percent for HIV infections with 18 of MDCs in the 5-20 percent range. An increase in the c-statistic of 5-20 percent does represent a substantial improvement. In general, the c-statistic tends to be very high in MDCs where death is relatively rare (e.g., for APR-DRGs the c-statistic is 0.9892 for the eye MDC) and lower in MDCs in which death is more frequent (e.g., for APR-DRGs the c-statistic is 0.7725 for the HIV MDC).

A composite measure of the magnitude of the difference in R^2 and c-statistic across the MDCs can be obtained by computing the weighted percent difference in R^2 and c-statistic across MDCs. Table 25 contains the weighted percent difference in R^2 and c-statistics relative to the Medicare DRGs. This is the same composite measure that was computed for cost in Table 5. The weighted percent difference in R^2 for APR-DRGs is much higher than the other DRG systems. For APR-DRGs the weighted percent difference in R^2 is 922.67, 3945.48 and 1974.03 percent for medical patients, surgical patients and all patients combined, respectively. The weighted percent difference in R^2 is very large, in part, because some high volume MDCs (e.g., MDC 15) have a very low R^2 for Medicare DRGs which result in a very large percent difference. The weighted percent difference in c-statistic for APR-DRGs is nearly twice as large as the other DRG systems. For APR-DRGs the weighted percent difference in c-statistic is 10.30, 18.03 and 11.07 percent for medical patients, surgical patients and all patients combined, respectively.

MORTALITY BY AGE AND SEX

The mortality rate can vary by the age and sex of the patient. The mortality rate predicted by a DRG system for patients with a specific age and sex can be computed by summing across all patients with the specified age and sex, the mortality rate (i.e., fraction died) in the DRG assigned to each patient and dividing the resulting sum by the number of patients with the specified age and sex. The predicted mortality rate can then be compared to the actual mortality rate. The predicted and actual mortality rate should be similar within each age and sex group. A large difference between predicted and actual mortality within an age and sex group would imply that the DRG system does not adequately account for the risk of mortality for patients with the age and sex characteristics. Table 26 shows the percent difference between the predicted and actual mortality rate for each of the DRG systems. The percent difference is com-

MDC	MDC Description	Medicare	SDRG		RDRG		AP-DRG		APR-DRG	
		R^2	R^2	% Diff	R^2	% Diff	R^2	% Diff	R^2	% Diff
1	Nervous	0.0631	0.1430	126.62	0.1267	100.79	0.1343	112.83	0.2513	298.25
2	Eye	0.0035	0.0016	54.28	0.0375	971.42	0.0232	562.85	0.1558	4351.42
3	Ear, Nose, Mouth and Throat	0.0773	0.1245	61.06	0.1408	82.14	0.0732	-5.30	0.2354	204.52
4	Respiratory	0.1173	0.1364	16.28	0.1447	23.35	0.1417	20.80	0.2195	87.12
5	Circulatory	0.0978	0.1923	96.62	0.1473	50.61	0.1351	38.13	0.2855	191.92
6	Digestive	0.0464	0.1174	153.01	0.1079	132.54	0.1005	116.59	0.2197	373.49
7	Hepatobiliary and Pancreas	0.1028	0.1704	65.75	0.1644	59.92	0.1371	33.36	0.2894	181.51
8	Musculoskeletal and Connective Tissue	0.0249	0.0938	276.70	0.0837	236.14	0.0702	181.92	0.2136	757.83
9	Skin, Subcutaneous Tissue and Breast	0.0705	0.0998	41.56	0.1001	41.98	0.0914	29.64	0.21.28	201.84
10	Endocrine, Nutritional and Metabolic	0.0267	0.0831	211.23	0.0718	168.91	0.0695	160.29	0.1723	545.31
11	Kidney and Urinary	0.0585	0.1101	88.20	0.0966	65.12	0.1019	74.18	0.1831	212.99
12	Male Reproductive	0.0851	0.1189	39.71	0.1154	35.60	0.0942	10.69	0.1853	117.74
13	Female Reproductive	0.1087	0.1506	38.54	0.1561	43.60	0.1100	1.19	0.2318	113.24
14	Pregnancy, Childbirth and the Puerperium	0.0006	0.0005	-16.66	0.0136	2166.66	0.0006	0.00	0.0717	11850.00
15	Newborns	0.0834	0.0830	-0.47	0.2956	254.43	0.3036	264.02	0.4155	398.20
16	Blood and Immunological	0.0062	0.0590	851.61	0.0664	970.96	0.0469	656.45	0.1806	2812.90
17	Myeloprofroliferative and Poorly Diff Neop	0.1400	0.2005	43.21	0.1533	9.50	0.1875	33.92	0.2588	84.85
18	Infectious and Parasitic	0.0657	0.1379	109.89	0.1079	64.23	0.1311	99.54	0.2374	261.33
19	Mental	0.0056	0.0157	180.35	0.0057	1.78	0.0051	-8.92	0.1533	2637.50
20	Alcohol and Drug	0.0051	0.0360	605.88	0.0009	-82.35	0.0044	-13.72	0.2015	3850.98
21	Injuries, Poisonings and Toxic Effects	0.0253	0.1044	312.64	0.0902	256.52	0.0574	126.87	0.2057	713.04
22	Burns	0.2346	0.2413	2.85	0.3263	39.08	0.2308	-1.61	0.4269	81.96
23	Factors Influencing Health Status	0.0089	0.0119	33.70	0.0331	271.91	0.0093	4.49	0.0923	937.07
24	Human Immunodeficiency Virus Infections	0.0053	0.0535	909.43	0.0530	900.00	0.0836	1477.35	0.1535	2796.22
25	Multiple Significant Trauma	0.0281	0.0283	0.71	0.0496	76.51	0.0713	153.73	0.1951	594.30
26	Other (Pre MDC)	0.1310	0.1651	26.03	0.1271	-2.97	0.1285	-1.90	0.2329	77.87

Table 23: R^2 for mortality by MDC

MDC	MDC Description	Medicare	SDRG		RDRG		AP-DRG		APR-DRG	
		C	C	% Diff	C	% Diff	C	% Diff	C	% Diff
1	Nervous	0.7758	0.8490	9.43	0.8377	7.97	0.8532	9.97	0.8989	15.86
2	Eye	0.7843	0.7484	-4.57	0.9620	22.65	0.9532	21.53	0.9892	26.12
3	Ear, Nose, Mouth and Throat	0.9263	0.9523	2.80	0.9668	4.37	0.9501	2.56	0.9814	5.94
4	Respiratory	0.8279	0.8464	2.23	0.8534	3.08	0.8593	3.79	0.8895	7.44
5	Circulatory	0.8154	0.8817	8.13	0.8595	5.40	0.8792	7.82	0.9146	12.16
6	Digestive	0.8130	0.8859	8.96	0.8688	6.86	0.8895	9.40	0.9261	13.91
7	Hepatobiliary and Pancreas	0.8622	0.9085	5.36	0.9046	4.91	0.9005	4.44	0.9406	9.08
8	Musculoskeletal and Connective Tissue	0.8353	0.9130	9.30	0.9199	10.12	0.9188	9.99	0.9592	14.83
9	Skin, Subcutaneous Tissue and Breast	0.8630	0.8904	3.17	0.9072	5.12	0.9192	6.51	0.9509	10.18
10	Endocrine, Nutritional and Metabolic	0.7532	0.8398	11.49	0.8090	7.40	0.8412	11.68	0.8832	17.25
11	Kidney and Urinary	0.8256	0.8780	6.34	0.8602	4.19	0.8763	6.14	0.9151	10.84
12	Male Reproductive	0.8682	0.8961	3.21	0.9508	9.51	0.9462	8.98	0.9681	11.50
13	Female Reproductive	0.9527	0.9687	1.67	0.9840	3.28	0.9763	2.47	0.9886	3.76
14	Pregnancy, Childbirth and the Puerperium	0.8015	0.7649	-4.56	0.8599	7.28	0.8096	1.01	0.8602	7.32
15	Newborns	0.9468	0.9454	-0.14	0.9651	1.93	0.9688	2.32	0.9729	2.75
16	Blood and Immunological	0.6427	0.7845	22.06	0.8023	24.83	0.7988	24.28	0.8901	38.49
17	Myeloprofroliferative and Poorly Diff Neop	0.8729	0.8956	2.60	0.8957	2.61	0.9069	3.89	0.9256	6.03
18	Infectious and Parasitic	0.7438	0.8294	11.50	0.8025	7.89	0.8242	10.80	0.8773	17.94
19	Mental	0.7805	0.7928	1.57	0.7889	1.07	0.7748	-0.73	0.9561	22.49
20	Alcohol and Drug	0.8584	0.8583	-0.01	0.6667	-22.33	0.8531	-0.61	0.9555	11.31
21	Injuries, Poisonings and Toxic Effects	0.7859	0.8851	12.62	0.8658	10.16	0.8824	12.27	0.9308	18.43
22	Burns	0.8272	0.8768	5.99	0.9603	16.09	0.8259	-0.15	0.9685	17.08
23	Factors Influencing Health Status	0.6970	0.7123	2.19	0.7834	12.39	0.6985	0.21	0.8479	21.64
24	Human Immunodeficiency Virus Infections	0.5508	0.6549	18.89	0.6805	23.54	0.6998	27.05	0.7725	40.25
25	Multiple Significant Trauma	0.6243	0.6234	-0.14	0.6846	9.65	0.7251	16.14	0.8035	28.70
26	Other (Pre MDC)	0.7791	0.8360	7.30	0.7745	-0.59	0.7763	-0.35	0.8752	12.33

Table 24: C-statistic for mortality by MDC

Statistic	Patients	SDRGs	RDRGs	AP-DRGs	APR-DRGs
R^2	All	98.30	391.14	94.95	1974.03
R^2	Surgical	302.36	968.27	311.12	3945.78
R^2	Medical	83.07	189.33	96.90	922.67
C	All	4.48	5.61	5.87	11.07
C	Surgical	7.79	12.49	11.16	18.03
C	Medical	4.29	4.69	5.20	10.30

Table 25: Weighted percent difference in R^2 and c-statistic across MDCs relative to Medicare DRGs

puted by subtracting the actual mortality rate from the predicted mortality rate and dividing by the actual mortality rate. The value of minus 15.7 for the Medicare DRGs for males age 0-28 days means that the predicted mortality rate is 15.7 percent lower than the actual mortality rate. The actual mortality displayed in Table 26 is for the Medicare DRGs. There were slight differences in the actual mortality across the different DRG systems due to differences in the number of patients in the excluded DRGs 469 and 470. The definition of chronic conditions for the pediatric patients is the same as the one used in the payment impact analysis. The predicted mortality rate for newborns and patients over age 79 was lower than the actual mortality rate while the predicted mortality rate for pediatric patients and adults under age 65 was higher than the actual mortality rate for both sexes and all DRG systems. For patients between ages 65 and 79, the predicted mortality was lower than the actual mortality rate for males and tended to be higher for females. The predicted and actual mortality rate differed the most for pediatric patients without chronic illnesses. The actual and predicted mortality rate for APR-DRGs were much more consistent

			Percent Difference Between Predicted and Actual Mortality Rate				
Sex	Age	Actual Mortality Rate	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs
Male	0-28 days	0.83	-15.7	-11.0	-18.5	2.6	-1.2
	29 days - 17 yrs wo chronic DX	0.31	145.2	131.3	138.7	193.3	48.4
	29 days - 17 yrs w chronic DX	0.98	77.6	72.6	87.2	117.9	10.7
	18-64 yrs	2.15	40.0	34.9	36.3	39.1	16.1
	65-79 yrs	5.39	-10.9	-7.8	-10.2	-13.2	-2.0
	80+ yrs	9.22	-37.0	-30.8	-33.8	-35.0	-21.1
Female	0-28 days	0.69	-17.4	-9.1	-13.0	10.6	7.7
	29 days - 17 yrs wo chronic DX	0.20	175.0	131.8	170.0	210.0	65.0
	29 days - 17 yrs w chronic DX	1.21	79.1	61.7	79.1	94.9	-3.4
	18-64 yrs	0.88	50.0	45.9	43.7	44.2	19.5
	65-79 yrs	4.22	0.9	1.2	1.4	-2.9	6.2
	80+ yrs	7.14	-26.9	-24.7	-26.5	-27.7	-14.6

Table 26: Percent difference between predicted mortality rate and actual mortality rate

than for the other DRG systems. These results show that none of the DRG systems fully account for the impact of age on patient mortality.

HOSPITAL LEVEL ANALYSIS FOR MORTALITY

The mortality analysis has focused on the ability of the DRG systems to predict mortality at the individual patient level. The DRGs can also be used to predict mortality at the hospital level. A regression was performed with the hospital actual mortality rate as the dependent variable and the predicted hospital mortality rate as the only independent variable. In addition, a multiple regression was performed using the same hospital characteristics that were used in the hospital level analysis for cost. The expectation is that the addition of hospital characteristics will not substantially increase the R^2 for mortality. Table 27 contains the R^2 results for mortality at the hospital level for each DRG system.

Relative to the Medicare DRGs, the hospital level R^2 is 4.84, 5.73, 4.87 and 9.92 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The addition of the hospital characteristics increased the R^2 for all the DRG systems. The percent increase in R^2 due to the addition of the hospital characteristics was 6.17, 5.31, 5.85, 5.28 and 3.60 for the Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The addition of the hospital characteristic did not substantially improve the hospital level R^2 for mortality for any of the DRG systems. This is in contrast to the hospital level R^2 for cost for which the addition of the hospital characteristics resulted in substantial improvement in R^2 for the Medicare DRGs, SDRGs and RDRGs.

SUMMARY AND CONCLUSIONS

The ability of the five DRG systems to predict cost and mortality were evaluated across a broad range of patient and hospital characteristics. The major findings from the analysis are as follows:

- There are substantial differences in the ability of the different DRG systems to predict patient cost. Relative to Medicare DRGs the untrimmed R^2 for cost is 8.1, 13.5, 15.0 and 30.3 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The percent increase in R^2 is greater for medical patients than for surgical patients.
- The primary source of the difference in R^2 for cost is the secondary diagnosis subgroups within each DRG system. The increase in untrimmed R^2 due to the secondary diagnosis subgroups ranges from 2.69 percent for the Medicare DRGs to 21.18 percent for APR-DRGs. The large contribution to R^2 due to the

	Medicare DRGs	SDRGs	RDRGs	AP-DRGs	APR-DRGs
R^2 expected mortality rate only	0.6139	0.6436	0.6491	0.6438	0.6748
R^2 expected mortality rate plus hospital characteristics	0.6518	0.6778	0.6871	0.6778	0.6991
Percent increase in R^2	6.17	5.31	5.85	5.28	3.60

Table 27: R^2 for mortality at hospital level across DRG systems

secondary diagnosis subclasses for APR-DRGs is likely the result of APR-DRGs being the only DRG system that explicitly recognizes the impact of multiple complications and comorbidities.

- The untrimmed R^2 for cost varies considerably across the individual MDCs ranging from 0.0417 for the mental health MDC for Medicare DRGs to 0.6274 for the newborn MDC for APR-DRGs.
- Within individual MDCs the untrimmed R^2 for cost differs substantially across the different DRG systems. For some MDCs the untrimmed R^2 for cost across the DRG systems differs by a factor of 3 or 4.
- The weighted percent difference in untrimmed R^2 for cost across MDCs relative to the Medicare DRGs is 18.78, 34.26, 30.45 and 70.62 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. Thus, at the MDC level the APR-DRGs have an untrimmed R^2 for cost that is seventy percent higher than the Medicare DRGs.
- Data trimming excludes more of the unexplained variability for DRG systems with a low untrimmed R^2. As a result of data trimming the R^2 for the different DRG systems is artificially compressed.
- There are substantial differences in which patients are trimmed in each of the different DRG systems. A disproportionate number of severity level 3 and 4 patients in APR-DRGs are high outliers in the other DRG systems and a disproportionate number of severity level 1 patients in APR-DRGs are low outliers in the other DRG systems.
- Relative to Medicare DRGs, the trimmed R^2 for cost is 3.5, 8.3, 8.7 and 16.7 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.
- The weighted percent difference in trimmed R^2 for cost across MDCs relative to the Medicare DRGs is 11.19, 20.62, 18.54 and 41.15 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.
- In general, the trimmed R^2 results for cost parallel the untrimmed results except that the magnitude of the differences are compressed.
- The R^2 for cost at the hospital level are 20 to 30 percent higher than at the patient level. Relative to the Medicare DRGs, the hospital level R^2 for cost is 1.27, 6.71, 14.63 and 21.15 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The addition of hospital characteristics increases the hospital level R^2 by 19.54, 20.67, 15.33, 7.98 and 4.83 for Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. With AP-DRGs and APR-DRGs the need for payment adjustments based on hospital characteristics is greatly reduced.
- In general, the more refined DRG systems, especially the APR-DRGs result in payment levels closer to hospital cost.
- The more refined DRG systems, especially APR-DRGs, tend to increase payments to non-normal neonates, pediatric patients with chronic conditions, patients 65 years and older, patients transferred in from acute care hospitals or SNFs, patients discharged to nonacute institutions, patients who died, Medicare, Medicaid and Title V patients and patients treated in children's hospitals and major teaching hospitals.
- Conversely, the more refined DRG systems, especially APR-DRGs, tend to decrease payments for normal newborns, pediatric patients without chronic conditions, adult patients age 18 to 64 years, patients discharged home, patients who left against medical advice,

commercial, self pay, workman's compensation and no charge patients and patients treated in rural hospitals.

- There are substantial differences in the ability of the different DRG systems to predict patient mortality. Relative to the Medicare DRGs, the percent increase in R^2 for mortality is 54.5, 42.4, 40.1 and 145.2 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. Relative to the Medicare DRGs, the percent increase in the c-statistic is 3.4, 2.9, 3.7 and 6.2 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.
- Within individual MDCs, the R^2 and c-statistic for mortality differs substantially across different DRG systems. The weighted percent difference in R^2 for mortality across MDCs relative to the Medicare DRGs is 98.30, 391.14, 94.95 and 1974.03 percent for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The weighted percent difference in the c-statistic for mortality across MDCs relative to the Medicare DRGs is 4.48, 5.61, 5.87 and 11.07, respectively.
- Across age and sex groups, the actual and predicted mortality are substantially closer for the APR-DRGs than the other DRG systems.
- Relative to Medicare DRGs the hospital level R^2 for mortality is 4.83, 5.73, 4.86 and 9.92 percent higher for SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively. The addition of hospital characteristics increases the hospital level R^2 for mortality by 6.17, 5.31, 5.85, 5.28 and 3.60 percent for the Medicare DRGs, SDRGs, RDRGs, AP-DRGs and APR-DRGs, respectively.

With the expansion of managed care, health plans are directing patients to selected hospitals which may appear more price competitive. States are increasingly disseminating information on hospital price and outcomes to consumers. A patient classification system is a critical component of these efforts and the various DRG systems are being widely used. Based on the results of the analysis there are several conclusions.

- If the DRGs are being applied to non Medicare patients then the Medicare DRGs, SDRGs and RDRGs are not appropriate. This is particularly true for newborn and pediatric patients with chronic diseases.
- If the DRGs are being used for payment purposes, the use of the more refined DRG systems, especially the APR-DRGs, greatly reduces the need for payment adjustments based on hospital characteristics.
- Children's hospitals and major teaching hospitals will be financially disadvantaged in price negotiations or a DRG based payment system unless there is a recognition of patient severity as provided in the more refined DRG systems.
- The only DRG system that is applicable to mortality is the APR-DRGs. In the other DRG systems, death is one of the variables used to define the DRGs. Even if death is removed as part of the definition of the DRGs, the ability of these DRG systems to predict patient mortality is substantially lower than the APR-DRGs, especially for certain age and sex groups.
- If the DRGs are being used for internal management, the severity levels in the more refined DRG systems provide a more clinically meaningful basis for communication of results to medical staffs.

The availability of severity adjusted pricing data and risk adjusted data on the quality of the services being purchased is a fundamental requirement for managed care programs to be effective. There are more than twenty states that publicly dis-

seminate information on provider performance. For example, the state of Florida annually publishes comparative information on charges, length of stay and mortality for all hospitals in the state. (State of Florida Agency for Health Care Administration, 1996) The Florida reports use APR-DRGs to adjust for severity of illness and risk of mortality.

The publicly disseminated information on provider performance will ultimately expand to include complication rates, readmission rates and other measures of efficiency and quality. In 1996, California implemented a requirement that hospitals report for each secondary diagnosis an indication of whether the secondary diagnosis was present at admission. (California Assembly Bill 3639, 1994) This requirement will allow meaningful complication rates to be computed, admission severity of illness and risk of mortality to be determined and patient deterioration or improvement during the hospitalization to be measured. For such data to be useful, it must be risk adjusted. If a multitude of risk adjustment methodologies are necessary to use such data then this will be a burden for providers and will make communication of results difficult. The APR-DRGs added risk of mortality subclasses to the basic DRG structure. Similarly, the DRG approach can be expanded to encompass other measures of resource use and quality.

The limiting factor for expanding the applications of DRGs is not in the DRG methodology, but the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) which is used to report diagnoses and procedures. With ICD-9-CM and the upcoming tenth revision of ICD many important patient characteristics can not be coded. In recent years there have been some improvements to ICD-9-CM. For example, a specification of birthweight was added as a fifth digit for the newborn codes and the number of days on mechanical ventilation was added to the mechanical ventilation codes. While these additions provide important information, there are still many clinical areas in which additional information is needed. For the evaluation of certain diseases, specific vital signs, laboratory results or health status information is essential. For example, the hematocrit would be useful in assessing patients with anemia or gastrointestinal bleeding. In a manner similar to the birth weight ranges for newborn codes a fifth digit could be added to the codes for these diseases which specifies the range of the hematocrit. In addition, functional health status measures such as activities of daily living provide important information. Patient frailty and limitations in mobility impact both mortality and the ability to recover. The additional information would only be collected on patients with specific diseases which minimizes the data collection burden on providers. The collection of a large number of patient characteristics across all patients is neither practical nor necessary. Using ICD-9-CM to obtain selective additional information when necessary is a practical and cost effective means of expanding the patient characteristics available to the DRGs. There are existing proposals on how ICD-9-CM could be expanded (Goldfield, et al., 1996).

With the appropriate additional information the DRG methodology can be adapted to apply to a wide spectrum of resource use and quality measures. The DRGs are the most widely used and well understood method of measuring casemix. As evidenced by the five DRG systems evaluated in this study, the DRG methodology has evolved and is continuing to evolve.

References

1. Averill, R, Goldfield, N, Steinbeck, B, Grant, T, *All Patient Refined Diagnosis Related Groups*, 3M Health Information Systems, Wallingford, Connecticut, 1995.
2. Averill, R, Steinbeck, B, Mullin, R, Goldfield, N, Elia, E, Anderman, S, Ray, K, Moore, S, Bielefeldt, R, *All Patient Diagnosis Related Groups Definitions Manual, Version 12*, 3M Health Information Systems, 1994.
3. Averill, R, Mullin, R, Steinbeck, B, Goldfield, N, Elia, E, *Diagnosis Related Groups Definitions Manual, Version 12.0*, 3M Health Information Systems, Wallingford, Connecticut, 1994.
4. California Assembly Bill 3639, Chapter 1063, Statues of State of California, 1994; effective January 1, 1996.
5. Federal Register, Vol. 16, No. 106, June 2, 1995, pg. 29209.
6. Feldman, R, "Are We Splitting Hairs over Splitting DRGs?", *Health Services Research,* Vol. 27, No. 5, December 1992, PP. 613-617.
7. Fetter, R, Shin, Y, Freeman, J, Averill, R, Thompson, J, "Case Mix Definition by Diagnosis-Related Groups", *Medical Care*, Vol. 18, No. 2, pp 1-53.
8. Freeman, J, Fetter, R, Hayong, P, Schneider, K, Lichenstein, J, Hughes, J, Bauman, W, Duncan, C, Freeman, D, Palmer, G, "Diagnosis Related Group Refinement with Diagnosis- and Procedure-Specific Comorbidities and Complications", *Medical Care,* Vol. 33, No. 8, pp 806-827.
9. Goldfield, N, Averill, R, Grant, T, Wang, Y, Graziano, D, Fillmore, H, *Improving the Prediction of Outpatient Services Using Patient Characteristics: Results of a Six-Month Prospective Study*, Final Report, 3M Health Information Systems, Wallingford, Connecticut, June 30, 1996.
10. HCIA, Inc., Baltimore, Maryland provided the data for the study, 1995
11. State of Florida Agency for Health Care Administration, *1996 Guide to Hospitals in Florida*, Tallahassee, Florida, 1996.

© 3M 1998